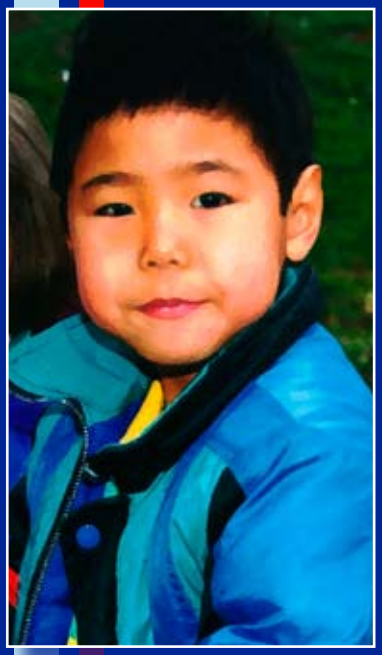
Evaluation of the North Carolina More at Four Pre-kindergarten Program

Year 5 Report

(July 1, 2005–June 30, 2006)



Children's Outcomes & Program Quality in the Fifth Year

February 2007

Ellen S. Peisner-Feinberg, Ph.D.
Jennifer M. Schaaf, Ph.D.
& the More at Four Evaluation Team

UNC
FPG Child Development Institute

We wish to acknowledge the members of our More at Four Evaluation Team who assisted with this phase of the research: Project coordinator (2005) Cathy Maris; Research assistants Lisa Hildebrandt, Diana Knechtel, Cyndee Lohr, and Yalitza Ramos; MAFKids Coordinator K.C. Elander: Programmers Steve Magers and Rita Slater; Statistician Lauren Nelson; and Data collectors Amber Alsobrooks, Becky Blissett, Vicki Boggs, Caroline Butler, Carolyn Cooper, Blair Edwards, Innana LaFevre, Eva Mamak, Kathleen MacGuire, Karen McGinley, Cathleen O'Donoghue, Mary Parrish, Teresa Partin, Kim Rangel, Jessica Rothstein, Cynny Scott, Mollie Stewart, Shannon Thorn, Hattie Warner, and Susan Wilson.

Special thanks to K.C. Elander for assistance with MAFPlan and MAFKids data and to Lisa Hildebrandt for publications assistance with this report.

Photographs: Don Trull, FPG Child Development Institute

In addition, we offer our appreciation to the teachers, administrators, and other staff of the More at Four programs across the state and to the families of More at Four children who provided these data.

Suggested citation: Peisner-Feinberg, E. S. & Schaaf, J.M. (2007). Evaluation of the North Carolina More at Four Pre-kindergarten Program: Children's Outcomes and Program Quality in the Fifth Year. Chapel Hill, NC: FPG Child Development Institute.

This research was funded by the North Carolina More at Four Pre-kindergarten Program, NC Office of School Readiness, NC Department of Public Instruction, as part of the statewide evaluation of the North Carolina More at Four Pre-kindergarten Program.

For more information about the Evaluation of the North Carolina More at Four Pre-kindergarten Program, visit the web site at www.fpg.unc.edu/~mafeval.

Table of Contents

List of Tables

List of Figures

Overview of the More at Four Program

The North Carolina More at Four Pre-kindergarten Program is a state-funded initiative for at-risk 4-year-olds, designed to help them be more successful when they enter elementary school. More at Four is based on the premise that all children can learn if given the opportunity, but at-risk children have not been given the same level of opportunity. The purpose of More at Four is to provide a high quality, comprehensive educational program for at-risk children during the year prior to kindergarten entry. The program first targets at-risk "unserved" children (those not already being served in a preschool program) and secondly, "underserved" children (those in a program but not receiving child care subsidies and/or those in lower quality settings). The More at Four Program was initiated in the 2001-2002 school year, with sites first serving children in the spring of 2002, and programs in all 100 counties since the 2003-2004 school year. More at Four served 17,251 children in the 2005-2006 school year, and has served over 49,000 children during the first five program years (2002-2006).

More at Four provides funding for classroom-based educational programs at a variety of sites designated by the local administration within each county or region (typically, either the local public school system or the local Smart Start partnership[a]). The programs are administered at the county or region (multi-county groupings) level with oversight by the State More at Four Office, and must include collaboration among the local school system(s), the local Smart Start partnership, and other interested members of the early childhood community (e.g., Head Start, child care providers, resource and referral agencies). Children are eligible for More at Four based on family income (up to 300% of Federal poverty status) and other risk factors (limited English proficiency, identified disability, chronic health condition, and developmental/ educational need). Priority for service is given first to at-risk children who are unserved in a preschool program at the time of enrollment, and second, to children who are underserved at enrollment. More at Four classrooms operate in a variety of settings, including public schools, Head Start, and community child care centers (both for-profit and nonprofit). Children may be enrolled in classrooms serving More at Four children exclusively or in blended classrooms also serving children funded through other sources such as Head Start or parent fees. The programs operate on a school-day and school calendar basis for 6 to 6-1/2 hours/day and 180 days/year. Local sites are expected to meet a variety of program guidelines and standards around curriculum, training and education levels for teachers and administrators, class size and student-teacher ratios, North Carolina child care licensing levels, and provision of other program services[1].

[a] Smart Start is a comprehensive early childhood initiative created in 1993 to ensure that all North Carolina children enter school healthy and ready to succeed. The program focuses on improving the quality of child care and providing health and family support services to children from birth to age five and their families. Program funds are distributed to 81 community partnerships serving all 100 North Carolina counties. For more information about Smart Start, visit the North Carolina Partnership for Children's website at http://www.ncsmartstart.org/.

Overview of the More at Four Evaluation

Since its inception in 2002, the statewide evaluation of the More at Four Program has been conducted by the FPG Child Development Institute at the University of North Carolina-Chapel Hill. The current report describes findings on the quality of the program and the outcomes for children during pre-k from two program years: 2003-2004 (year 3) and 2005-2006 (year 5). Separate results for the 2005-2006 cohort are presented to provide a picture of the program in its most recent year of operation as well as results from the 2 cohorts combined to examine the robustness of these findings over time. Previous reports are available with detailed results for the 2003-2004 year[2], as well as results for the pre-k program in 2001-2002[3], 2002-2003[4], and 2004-2005[5], and results from a kindergarten follow-up of children in 2004-2005[6].

The primary research questions addressed by this evaluation included:

- What were the characteristics of the local programs?
- What was the quality of the services provided?
- What were the outcomes of children attending the More at Four Program?
- What factors were associated with better outcomes for children?

In order to address these questions, we gathered information from three sources: monthly service reports, observations of classroom quality, and individual child assessments. The monthly service report data from each local contractor provided information about characteristics of the program and the children served for all sites, classrooms, and children participating in More at Four in their county or region (multi-county group). Observations were conducted in randomly-sampled classrooms to provide information about the quality of classroom practices, including the activities and materials provided, the interactions among teachers and children, the physical environment, and the daily organization of the program (n=99 classrooms in 2003-2004, n=57 classrooms in 2005-2006). In addition, information about the quality of the classroom literacy environment and the sensitivity of teacher-child interactions was gathered in 2005-2006. Individual assessments of children's skills in these randomly-selected classrooms were conducted near the beginning and end of the program year to provide information about their developmental growth and school readiness over the pre-k year (n=514 children in 58 classrooms in 2003-2004, n= 478 children in 57 classrooms in 2005-2006). These assessments included measures of children's language and literacy skills, math skills, general knowledge, and social skills, in accord with generally accepted definitions of school readiness, including the recommendations of the National Education Goals Panel.[7]

Methods

CLASSROOM QUALITY OBSERVATION METHODS

Classroom quality was examined in two samples of More at Four pre-kindergarten classrooms. The first sample included More at Four classrooms operating in 2003-2004 and the second sample included classrooms operating in 2005-2006.

Participants

Classroom observations in both years were conducted in a sample of More at Four classrooms randomly selected from those that began serving children by the beginning of September of the study year to insure that children had the opportunity for a full program year. In 2003-2004, observations were conducted in 99 classrooms, including 57 of the 58 classrooms from which the child sample described subsequently was drawn (one of the 58 classrooms was no longer part of the More at Four program at the time of the classroom observations). In 2005-2006, observations were conducted in the same 57 classrooms from which the child sample was drawn. These included 2 first-year classrooms, 8 second-year classrooms, 15 third-year classrooms, 18 fourth-year classrooms, and 14 fifth-year classrooms. The classroom year of operation was not available for the 2003-2004 sample.

Procedures

Observations of classroom quality were conducted in the spring of each school year. The first sample was collected in 2004 (3/19/04-6/3/04) and the second in 2006 (3/9/06-5/2/06). Observations typically lasted 4 to 5 hours per classroom. Data collectors were trained to an acceptable criterion of reliability prior to gathering data. Interrater reliability data were collected in the field for 20% of each sample. Reliability data for the ECERS-R yielded a kappa of .85 for the 2003-2004 sample and a kappa of .74 for the 2005-2006 sample. Reliability data for the CIS (2005-2006 only) yielded a kappa of .77. For the ELLCO (2005-2006 only), reliability data yielded a kappa of .53 for the Classroom Observation Scale, and exact agreement scores of 84% on the Literacy Environment Checklist and 86% on the Literacy Activities Rating Scale.

Measures

Global classroom quality was assessed in both samples using the Early Childhood Environment Rating Scale-Revised[8] (ECERS-R), an observational rating scale that measures the developmental appropriateness of classroom practices including the activities and materials provided, the interactions among teachers and children, the physical environment, and the daily organization of the program. The scale contains 43 items arranged into 7 subscales: Space and furnishings, Personal care routines, Language-reasoning, Activities, Interaction, Program structure, and Parents and staff. Each subscale item is rated on a 7-point scale from low to high, where 1 = "inadequate," 3 = "minimal," 5 = "good," and 7 = "excellent".[a] In the current study, the total and subscale scores were computed as mean item scores ranging from 1.0 to 7.0, with higher scores indicating better classroom quality. The ECERS-R and its predecessor, the ECERS, have been used in a wide range of early education research studies. The scales have

[a] Program guidelines for More at Four state that participating classrooms should score at least 4.5 on this rating scale by the second year of operation.

been demonstrated to have good interrater reliability (total scale r = .92)[8] and predictive validity[9].

In the 2005-2006 classroom sample, two additional measures were completed for the participating classrooms: the Early Language and Literacy Classroom Observation[10] (ELLCO) and the Caregiver Interaction Scale[11] (CIS). The ELLCO measures the extent to which classrooms provide children optimal support for language and literacy development. This observational rating scale includes three scales: Classroom Observation Scale, Literacy Environment Checklist, and Literacy Activities Rating Scale, each scored on a different metric. The Classroom Observation Scale consists of 14 items across 2 subscales: General classroom environment and Language, literacy, and curriculum. Each item is scored on a 1-5 scale, where 1 = "deficient", 3 = "basic", and 5 = "exemplary". Mean item scores, ranging from 1.0-5.0, were used in the present study. The Literacy Environment Checklist has a total score ranging from 0-41, based on 5 subscales: Book area (0-3), Book selection (0-8), Book use (0-9), Writing materials (0-8), and Writing around the room (0-13). The Literacy Activities Rating Scale has a total score ranging from 0-13 and contains two subscales: Reading (0-8) and Writing (0-5). These scales have demonstrated good interrater reliability (Classroom Observation Scale=90%, Literacy Environment Checklist=88% within 1 point, and Literacy Activities Rating Scale=81%) and moderate to good internal consistency (Cronbach's alpha: Classroom Observation Scale=.90, Literacy Environment Checklist=.84, Literacy Activities Rating Scale=.66)[10].

The CIS measures the sensitivity of teacher's interactions with children. It includes 26 items divided into 4 subscales: Sensitivity, Harshness, Detachment, and Permissiveness. Each item is scored on a 1-4 scale from "not at all" to "very much". Mean item scores ranging from 1.0 to 4.0 were calculated for each subscale. For the total score, scores on the three negative subscales (Harshness, Detachment, and Permissiveness) were reversed and a total mean item score was calculated where higher scores indicated more positive teacher-child interactions. The scale has demonstrated good interrater reliability of 80%[11].

CHILD OUTCOMES ASSESSMENT METHODS

Two samples of children participating in the More at Four Program were included for the present report. The first sample participated in the More at Four Program in 2003-2004 and the second in 2005-2006. Individual assessments of children's language and literacy skills, math skills, general knowledge, and social skills were conducted near the beginning and end of each program year to provide information about children's growth over the program year.

Participants

Children were recruited from randomly-selected More at Four classrooms across North Carolina. These classrooms also participated in observations of classroom quality described previously. The first sample included 58 classrooms and 514 children in fall 2003 and 434 of the same children in spring 2004. The second sample included 57 classrooms and 478 children in fall 2005 and 445 of the same children in spring 2006.

Sample Selection. In both years, a random sample of classrooms was selected from those that began serving children by the beginning of September of the study year in order to insure that children had the opportunity for a full program year.

For the 2003-2004 sample, 58 More at Four classrooms were randomly selected from 599 eligible classrooms. We attempted to recruit all More at Four children enrolled in the selected classrooms and obtained an overall consent rate of 85% (573/675). Children with parental consent who were absent or had withdrawn from the program at the time of data collection were not assessed, resulting in a sample of 514 children. Comparisons of assessed children to all other More at Four children indicated that the two groups were similar in terms of the distributions on most demographic characteristics, including age, gender, poverty status, risk factor total, limited English proficiency, disability, health condition, and family size. There were some differences in terms of ethnicity, service priority status, and attendance. The assessed group had fewer African-American and Asian children, a slightly higher average service priority level, and greater days of attendance. (See Table 1.)

For the 2005-2006 sample, 57 classrooms were included. Of these, 53 were randomly selected from 952 eligible classrooms and 4 additional classrooms also participating in the North Carolina Department of Public Instruction Model Literacy Program were added to the sample. We attempted to recruit all More at Four children from each classroom up to a maximum of 10. In cases where more than 10 More at Four children had parental consent, 10 children were randomly selected to participate. The overall consent rate was 81% (687/846), with a final sample of 478 children. Comparisons of assessed children to all other More at Four children indicated that the two groups were similar in terms of the distribution on most demographic characteristics, including age, gender, poverty status, risk factor total, identified disability, health condition, and family size. There were some differences in terms of ethnicity, English proficiency, service priority status, and attendance. The assessed group had fewer African-American and more Latino children, a higher proportion of children with limited English proficiency, a higher average service priority level, and greater days of attendance. (See Table 1.)

Table 1. Characteristics of Assessed and Non-Assessed Children

Factor[a]		2003-2004 N=10,891		2005-2006 N=17,251	
		Assessed (n=514)	**Non-Assessed** (n=10,377)	**Assessed** (n=478)	**Non-Assessed** (n=16,773)
Child age on 10/16[b] (Mean)		4.5	4.5	4.5	4.5
Gender[c] (% female)		50.4%	48.4%	49.8%	49.0%
Ethnicity (%)	Black/African-American	36.8%	43.1%*	30.1%	36.6%*
	White/European-American	36.4%	31.0%	32.6%	34.2%
	Hispanic/Latino	16.9%	17.8%	28.0%	21.7%*
	Other/Multiracial	9.1%	6.5%	8.0%	6.1%
	Asian	0.8%	1.7%*	1.3%	1.5%
Poverty Status (%)	Free Lunch Eligible	75.1%	74.7%	74.9%	73.6%
	Reduced Price Eligible	15.6%	14.9%	17.2%	16.3%
Risk Total[d] (Mean)		1.9	1.9	2.0	1.9
Individual Risk Factors (%)	Limited English Proficiency[e]	17.3%	18.0%	24.3%	18.4%*
	Identified Disability[f]	6.2%	7.1%	4.0%	4.8%
	Chronic Health Condition[g]	2.3%	3.4%	4.8%	3.8%
Service Priority Status[h] (Mean)		1.8	2.0**	2.2	2.7***
Total Days of Attendance (Mean)		149.1	123.5***	155.5	135.2***
Family Size[i] (Mean)		4.1	4.0	4.0	4.1

[a] Significant comparisons reported represent differences between the two groups based on T-tests or chi-square tests with a Bonferroni correction for multiple comparisons. Significance levels are *p< .05, **p< .01, ***p< .001.
[b] In Year 3, age was not reported for 1 child
[c] In Year 3, gender was not reported for 49 children.
[d] In Year 3, risk total was not available for 58 children.
[e] In Year 3, limited English proficiency was not reported for 58 children.
[f] In Year 3, identified disability was not reported for 58 children.
[g] In Year 3, health condition was not reported for 58 children.
[h] The categories for service priority status levels changed from 2003-2004 to 2005-2006, from 5 levels to 8 levels. Note that lower values represent higher service priority.
[i] In Year 3, family size was not reported for 15 children.

Child Characteristics. In the 2003-2004 sample, the average child age was 4.5 years (range = 4.0-5.0 years) at the time of the fall assessments and 5.1 years (range 4.6–5.6 years) at the time of the spring assessments. At the time of study enrollment, half (50%) of the children were female and half were male; 37% were African-American, 36% Caucasian, 17% Latino, and 10% were from other ethnic/racial or multiracial groups.

In the 2005-2006 sample, the average child age was 4.5 years (range = 4.0-5.1) at the time of the fall assessments and 5.1 years (range=4.5-5.6) at the time of the spring assessments. At the time of study enrollment, half (50%) of the children were female and half were male; 30% were African-American, 33% Caucasian, 28% Latino, and 8% were from other ethnic/racial or multiracial groups.

Procedures

Two sources of child outcomes data were gathered: Individual assessments of children's language and cognitive skills and teacher ratings of children's behavioral skills. Assessment data from the first sample were collected in fall 2003 (9/20/03-11/7/03) and again in spring 2004 (4/28/04-6/10/04). Assessment data from the second sample were collected in fall 2005 (9/22/05-11/22/05) and spring 2006 (4/26/06-6/8/06). Child assessments were conducted on-site at each school or child care center by trained data collectors, and lead teachers were asked to complete rating scales following the assessments.

Measures

The child assessment battery consisted of eight measures focusing on language and literacy skills, pre-math skills, and general knowledge. Lead teachers also rated each child's social skills and problem behaviors in the classroom. The outcome areas measured were consistent with the recommendations of the National Education Goals Panel[7] for defining school readiness. (See Table 2 for an overview of these measures.) In addition, children were administered three subscales of the PreLAS 2000[12] (Simon Says, Art Show, and The Human Body), an individual assessment designed to measure young children's oral language proficiency in English, including both receptive and expressive language ability. This measure was used to adjust for children's English language proficiency in the analyses, as well as to examine English language proficiency as a moderator of program effects. Fluency scores ranging from 1-5 were calculated, where 1=Non-English speaker, 2-3=Limited English speaker, and 4-5=Fluent English speaker.

All children were administered the eight child assessment measures plus the English proficiency assessment in both the fall and spring. In the 2005-2006 sample, children who spoke Spanish and scored below the fluent level on the PreLAS 2000 in the fall were also administered the same measures in Spanish in separate sessions in both the fall and spring. It is important to note that for the standardized measures (receptive language, rhyming, applied problems), the English and Spanish versions differed somewhat in content, while for the remaining measures, the items on the English and Spanish versions were direct translations of one another. A total of 120 children were assessed in both languages.

Table 2. Child Outcome Measures

Domain	Measure	Skills Assessed	Scoring
Language and literacy	Peabody Picture Vocabulary Test-III (PPVT-III)[13] Test de Vocabulario en Imagenes Peabody (TVIP)[14]	Receptive vocabulary	Standardized measure, Mean=100, SD=15
	Woodcock Johnson-III Tests of Achievement (WJ-III)[15] Rhyming (subtest 21A, Sound Awareness test) Batería III Pruebas de aprovechamiento[16] Rima (Prueba 21A, Discernimiento de sonidos)	Phonological awareness	Range=0-17
	Naming Letters Task[17] (English and Spanish versions)	Alphabet knowledge	Range=0-26
	Story and Print Concepts Task[18] (English and Spanish versions)	Early literacy skills including book knowledge, story comprehension, and print awareness	Range=0-14
Math	Woodcock Johnson-III Tests of Achievement[15] Applied Problems Test (Test 10) Batería III Pruebas de aprovechamiento[16] Problemas aplicados (Prueba 10)	Ability to solve practical math problems including counting, simple addition, and subtraction	Standardized measure, Mean=100, SD=15
	Counting Bears Task[19] (English and Spanish versions)	Ability to count in one-to-one correspondence	Range=0-40
General knowledge	Social Awareness Task[20] (English and Spanish versions)	Knowledge of child's full name, age and birth date	Range=0-6
	Color Bears Task[21] (English and Spanish versions)	Knowledge of 10 basic colors	Range=0-20
Classroom behavior	Social Skills Rating System (SSRS) Social Skills subscale[22]	Social skills (e.g., "follows your directions")	Standardized measure, Mean=100, SD=15
	Social Skills Rating System (SSRS) Problem Behaviors subscale[22]	Problem behaviors (e.g., "argues with others")	Standardized measure, Mean=100, SD=15

Program Characteristics

Information about the characteristics of the More at Four Program, including the local sites, the classrooms, and the children served in the 2003-2004 and 2005-2006 program years, is described below, to give a picture of program operations over this period of interest. In most cases, the characteristics of the local More at Four sites have remained fairly similar over time.

The More at Four Program has grown substantially since its inception in the 2001-2002 school year. The program served 1,244 children during that first year, compared to 10,891 in 2003-2004 (year 3) and 17,251 in 2005-2006 (year 5). Local sites have existed in all 100 counties (91 local contractors) since the 2003-2004 school year. The mean More at Four class size was 16 each year, with a median of 18 (the maximum allowable based on program guidelines). The median proportion of children in each classroom participating in More at Four was three-quarters or more each year, indicating that the majority of children in the classrooms were enrolled in this program. The program has targeted "unserved' children, with the majority of children classified as unserved at the time of enrollment (83% in 2003-2004 and 79% in 2005-2006). (See Table 3 for more information about various program characteristics for the 2003-2004 and 2005-2006 program years.)

As shown in Figure 1, nearly half (45% in 2003-2004 and 47% in 2005-2006) of the children were served in public preschool sites and almost one-third (32% and 29%, respectively) were served in private for-profit child care settings. Smaller proportions were served in private nonprofit child care settings (11% and 12%, respectively) and Head Start sites, including those administered by public schools (10% and 12%, respectively).

The characteristics of the More at Four classrooms have remained fairly similar over time as well. More at Four program guidelines recommend that classrooms use a research-based curriculum. As seen in Table 4, more than three-quarters of the classrooms each year reported using Creative Curriculum[23] as their primary curriculum, with smaller numbers reporting using OWL[24] or Bright Beginnings[25] (14% and 15%)[a], High/Scope[26] (8% and 7%), or Montessori (less than 1%).

One area that has shown some change over this period is teacher qualifications, which have indicated some improvement. The program guidelines require that lead teachers have a B-K license (or the equivalent) within four years. As seen in Table 5, the percentage of teachers with Bachelor's degrees or higher has risen slightly in public school settings while the percentage with High School degrees only has decreased across all settings. Similarly, as shown in Table 6, the percentage of teachers with a B-K license (or equivalent) has increased by more than 25% (from 39% to 49%) over this time period, although the change is primarily occurring in public school settings. In contrast, in community settings, the percentage of teachers with some type of early childhood credential (CDA or NCECC) has increased substantially (from 20% to 38%) while the percentage with no credential has decreased (from 53% to 37%).

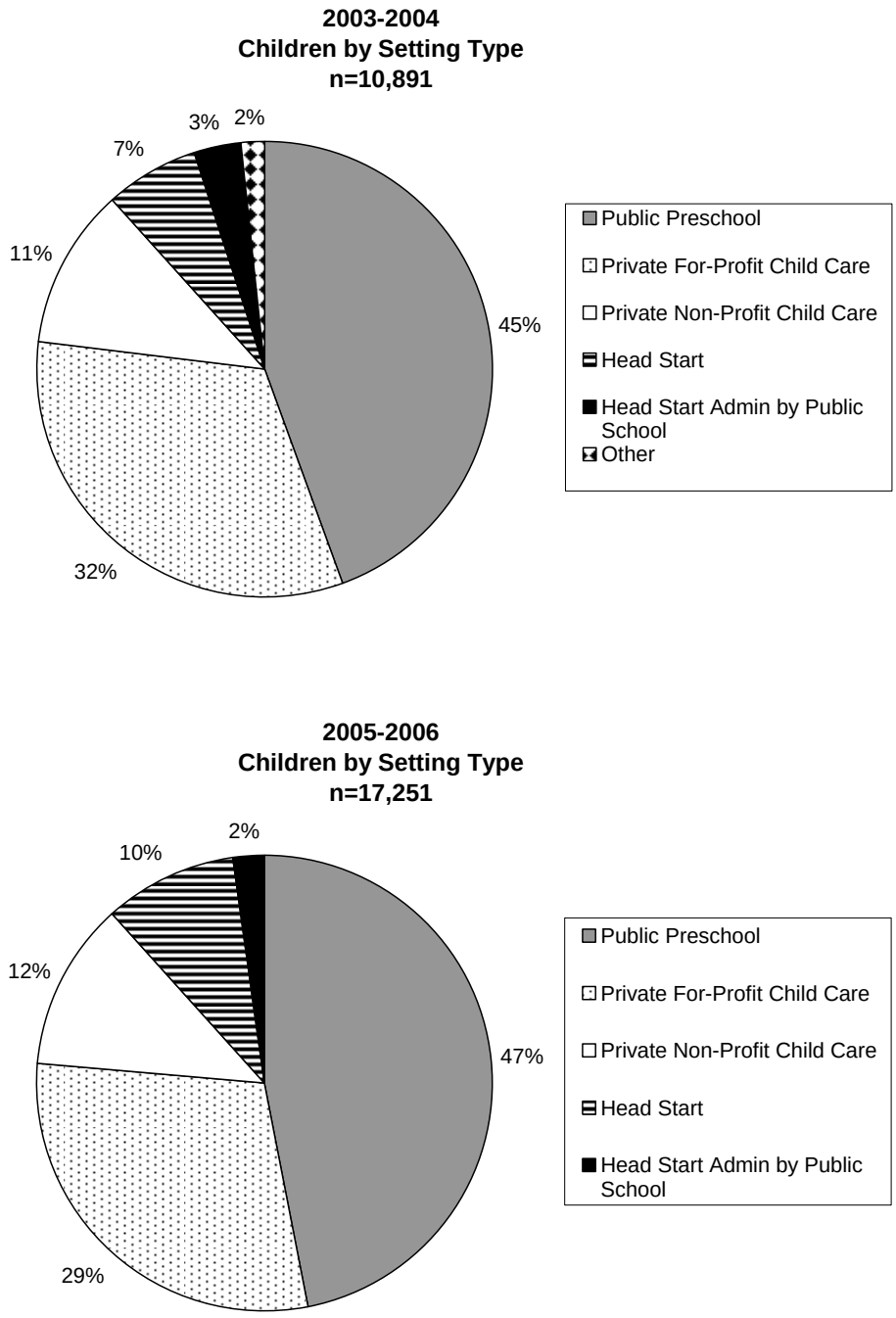
[a] The Bright Beginnings curriculum was changed to the OWL curriculum (Opening the World of Learning) in the 2004 edition.

Table 3. Program Characteristics

Program Characteristic	Year 3 2003-2004	Year 5 2005-2006
Total More at Four Local Contractors	91	91
Total More at Four Counties	100	100
Total More at Four Sites (Facilities)	628	790
Total More at Four Classrooms	883	1,218
Total Children Served	10,891	17,251
Total Children Not Served at Time of Enrollment[a]	9,070 (83%)	13,617 (79%)
Total Children Never Previously Served[a]	6,788 (62%)	10,325 (60%)
Average Class Size[b]		
Mean	16.3	16.2
Median	17.6	17.6
SD	2.6	2.7
Average Number of More at Four Children per Class[c]		
Mean	10.7	12.3
Median	10.6	13.6
SD	5.8	4.9
Average Proportion of More at Four Children per Class[d]		
Mean	0.67	0.76
Median	0.78	0.91
SD	0.32	0.26

[a] These data are based on reported service priority status.
[b] These data are based on the monthly reported total class size, including both More at Four and non-More at Four children. The More at Four program guidelines indicate a maximum class size of 18. In 2003-04, 2 classrooms were granted exceptions to exceed the maximum allowable class size, with an average class size of 20 and 23 children. In 2005-2006, 52 classrooms exceeded the maximum allowable class size with an average class size of 19; 50 of these classrooms were located in one contract that was granted an exception to the maximum class size.
[c] These data are based on the monthly reported number of More at Four children for each classroom.
[d] These data are based on the proportion of the monthly reported number of More at Four children and class size for each classroom.

Figure 1. Distribution of Children by Setting Type by Year[a]



[a] Children who attended more than one More at Four site (in 2003-2004, 230 children attended 2 or more sites and in 2005-2006, 331 children attended 2 or more sites) are represented by the setting type in which they were enrolled the longest.

Table 4. Distribution of Classrooms by Primary Curriculum Type

Curriculum Type[a]	**Year 3 2003-2004** n=871[b]	**Year 5 2005-2006** n=1,218
Creative Curriculum	76.5% (666)	77.9% (949)
Bright Beginnings/ OWL[c]	13.9% (121)	14.7% (179)
High/Scope	7.7% (67)	6.7% (82)
Montessori	0.5% (4)	0.3% (4)
Other	1.5% (13)	0.3% (4)

[a] The Bank Street curriculum was also included in the guideline recommendations, but no classrooms reported it as the primary curriculum.
[b] In Year 3, curriculum was not reported for 12 classrooms.
[c] The Bright Beginnings curriculum was changed to the OWL curriculum (Opening the World of Learning) in the 2004 edition. In 2005-2006, 132 (10.8%) of the programs reported using Bright Beginnings and 47 (3.9%) reported using OWL.

Table 5. Education Levels of More at Four Lead Teachers

Highest Degree Earned	2003-2004			2005-2006		
	Public School Settings n=449[a]	**Community Settings** n= 535[b]	**All Settings** n=984	**Public School Settings** n=725	**Community Settings** n=617	**All Settings** n=1342
MA/MS or higher	17.2% (77)	4.1% (22)	10.1% (99)	13.8% (100)	3.4% (21)	9.0% (121)
BA/BS	77.1% (346)	62.6% (335)	69.2% (681)	84.6% (613)	60.9% (376)	73.7% (989)
AA/AAS	2.5% (11)	25.2% (135)	14.8% (146)	1.4% (10)	31.8% (196)	15.4% (206)
HS diploma/ GED	3.3% (15)	8.0% (43)	5.9% (58)	0.3% (2)	3.9% 24	1.9% (26)

[a] These data were not reported for 4 public school lead teachers.
[b] These data were not reported for 1 community setting lead teacher.

Table 6. Licensure/Credential Levels of More at Four Lead Teachers

Highest License/ Credential[a]	2003-2004			2005-2006		
	Public School Settings n=453	**Community Settings** n=536	**All Settings** n=989	**Public School Settings** n=725	**Community Settings** n=617	**All Settings** n=1342
B-K or Preschool add-on License	66.2% (300)	15.9% (85)	38.9% (385)	77.8% (564)	15.4% (95)	49.1% (659)
Provisional B-K License	1.8% (8)	0.8% (4)	1.2% (12)	5.1% (37)	1.1% (7)	3.3% (44)
Other Teacher's License	18.3% (83)	10.4% (56)	14.1% (139)	9.8% (71)	8.6% (53)	9.2% (124)
CDA Credential	0% (0)	3.9% (21)	2.1% (21)	0.5% (4)	6.5% (40)	3.3% (44)
NCECC	1.1% (5)	16.2% (87)	9.3% (92)	1.1% (8)	31.4% (194)	15.1% (202)
None	12.6% (57)	52.8% (283)	34.4% (340)	5.7% (41)	37.0% (228)	20.0% (269)

[a] Note: B-K = Birth-Kindergarten, CDA = Child Development Associate, NCECC = North Carolina Early Childhood Credential. Other teacher's license includes non-early childhood licenses and licenses from other states.

The population of children participating in More at Four has continued to be at-risk and of high service priority status, as intended. The demographic characteristics of the children served have remained fairly constant over time. Approximately half the children served are boys and half are girls. The percentage of Latino children has increased slightly while the percentage of African-American children has decreased. Children live in households with an average of 4 members and most (69%-79%) of the children's primary caregivers were employed. (See Table 7.) In terms of risk status, the programs continued to serve a primarily low-income population, with almost three-quarters of the children living in families at or below 130% of poverty (i.e., eligible for free lunch) and another 15%-16% between 131%-185% of poverty (i.e., eligible for reduced-price lunch). Almost one-fifth of the children served spoke limited or no English at entry into the program. One shift has been seen in the percentage of children with an identified disability, with a decrease from 7% to less than 5% over this two-year period. (See Table 8.) The majority of children participating in More at Four were unserved (i.e., not being served in a child care or preschool program) at the time of entry into the program (83% in 2003-2004 and 79% in 2005-2006), the primary target group for the program. (See Table 9.)

Table 7. Characteristics of More at Four Children

Characteristic		**2003-2004 n=10,891[a]**	**2005-2006 n=17,251[b]**
Gender	Male	51.5% (5,588)	51.0% (8,803)
	Female	48.5% (5,254)	49.0% (8,448)
Ethnicity	Black/African American	42.8% (4,658)	36.4% (6,277)
	White/European American	31.3% (3,404)	34.1% (5,890)
	Hispanic/Latino	17.8% (1,934)	21.8% (3,765)
	Multiracial	3.4% (369)	3.5% (604)
	Native American/Alaskan Native	3.0% (328)	2.4% (407)
	Asian	1.6% (176)	1.5% (263)
	Native Hawaiian/Pacific Islander	0.2% (22)	0.3% (45)
Median Total Household Size		4	4
Percent of Primary Caregivers Employed		69.2% (7,535)	79.3% (13,385)

[a] In Year 3, gender was not reported for 49 children, household size was not reported for 105 families and primary caregiver's employment was not reported for 14 families.
[b] In Year 5, primary caregiver's employment was not reported for 369 families.

Table 8. Risk Factor Status of Children

Risk Factor[a]	**Risk Factor Level**	**Year 3 2003-2004** n=10,833[b]	**Year 5 2005-2006** n=17,251
Family Income[c]	Below 130% of poverty (eligible for free lunch)	74.3% (8,051)	73.6% (12,694)
	131-185% of poverty (eligible for reduced price lunch)	15.3% (1,653)	16.4% (2,820)
	186-200%	10.4% (1,129)	3.6% (615)
	201-250%		4.8% (827)
	251-300%		1.7% (295)
Limited English Proficiency	Family and/or child speak limited or no English in the home	18.1% (1,958)	18.6% (3,209)
Developmental/ Educational Need[d]	Developmental/educational need indicated by performance on a developmental screen	---	15.6% (2,694)
Identified Disability	Child has an IEP	7.0% (762)	4.8% (831)
Chronic Health Condition(s)	Child is chronically ill/ medically fragile	3.3% (361)	4.7% (818)

[a] In Year 3, sites could choose to use either Model I or Model II guidelines for determining risk levels; 75% used Model I and 25% used Model II. Only Model I was available in previous years and only Model II was available in subsequent years. For more information, see the Year 3 evaluation report.[3]
[b] In Year 3, risk factor data were not reported for 58 children.
[c] In Year 3, only one category for family income levels at or above 186% of poverty was distinguished under Model I.
[d] In Year 3, developmental/educational need was an additional risk factor only for Model II guidelines and only for children whose family incomes were 251-300% of poverty. In 2003-2004, 6 children in this category were identified as having a developmental/educational need. In 2005-2006, developmental/educational need was included as a risk factor for children in all income categories.

Table 9. Distribution of Children by More at Four Service Priority Status at Time of Enrollment

Service Priority Status[a]	**Year 3 2003-2004** n=10,891	**Year 5 2005-2006** n=17,251
Unserved		
Children who have never been served in any preschool or child care setting.	62.3% (6,788)	59.9% (10,325)
Children who are currently unserved (may previously have been in child care or preschool program) and are on the subsidy waiting list.	9.8% (1,072)	5.5% (957)
Children who are currently unserved (may previously have been in child care or some other preschool program) and are not eligible for subsidy.	11.1% (1,210)	7.6% (1,313)
Children who are in a child care situation and served for 5 months or less in the year prior to More at Four.	--[b]	5.9% (1,022)
Underserved		
Children who are eligible for subsidy but are not receiving it (but are in some kind of child care or preschool program).	5.6% (606)	2.1% (364)
Children who are in unregulated child care that does not meet the More at Four Pre-K standards.	--[b]	4.2% (716)
Other children, including those in pre-kindergarten or child care that does not meet More at Four standards.	11.2% (1,215)	7.2% (1,236)
Children served by this site as 3-year-olds.	--[b]	7.6% (1,318)

[a] Note that all children served must first meet the eligibility requirements as defined in the More at Four Program Guidelines.
[b] The program guidelines for service priority status did not distinguish this category in Year 3.

Results

CLASSROOM QUALITY

In order to gather information about the quality of educational practices in the More at Four classrooms, observations were conducted in a random sample of classrooms. These included 57 classrooms in 2005-2006 and 99 classrooms in 2003-2004. Data were gathered about the developmental appropriateness of classroom practices using the ECERS-R[8], including the activities and materials provided, the interactions among teachers and children, the physical environment, and the daily organization of the program. In the 2005-2006 cohort, observational data were also gathered about the quality of the literacy environment of the classroom using the ELLCO[10] and the sensitivity of teacher-child interactions using the CIS[11].

Classroom Practices

The average scores on the ECERS-R (which is scored on a 1-7 scale from inadequate to excellent) for both the 2005-2006 and 2003-2004 cohorts are presented in Table 10. The scores for the 2005-2006 cohort were somewhat lower than in previous years[a]. The average total score for the 2005-2006 cohort was 4.4, which is in the medium quality range (i.e., scores between 3.0-4.9), compared to an average score for the 2003-2004 cohort of 5.3, in the high quality range (i.e., scores between 5.0-7.0). Detailed findings for the 2003-2004 cohort are presented in an earlier report[2], and are similar to previous years.

As seen in Figure 2, 12% of classrooms scored in the high quality range (5.0-7.0), 86% scored in the medium quality range (3.0-4.9), and 2% scored in the low quality range (1.0-2.9) in the later cohort, compared to 76% high, 24% medium, and 0% low, respectively, in the earlier cohort. In addition, 47% of the classrooms in 2005-2006 scored at or above 4.5 (as required by the More at Four program guidelines), compared to 88% in the earlier cohort.

As seen in Figure 3, the average subscale scores were also similarly lower for the later cohort, with most scores in the medium quality range (Space and furnishings, Language and reasoning, Activities, Interaction, and Program structure). One subscale scored in the high quality range (Parents and staff) and one in the low quality range (Personal care routines). While the pattern of scores was similar for both cohorts, for the earlier cohort, five subscale scores were in the high quality range (Space and furnishings, Language and reasoning, Interaction, Program structure, and Parents and staff), and the remaining two were in the medium quality range (Personal care routines and Activities).

[a] A separate assessment of a sample of 204 second-year classrooms by a different project in 2005-2006 found a mean score of 5.9 on the ECERS-R. The present study sample included 2 first-year classrooms, 8 second-year classrooms, 15 third-year classrooms, 18 fourth-year classrooms, and 14 fifth-year classrooms.

Table 10. Quality of Classroom Practices (ECERS-R)

Item Description	2003-2004 (n=99)			2005-2006 (n=57)		
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**
Total Score[a]	**5.3**	**0.6**	**3.4-6.4**	**4.4**	**0.7**	**2.8-5.8**
Total Child Items Score[b]	**5.3**	**0.7**	**3.0-6.6**	**4.2**	**0.7**	**2.7-5.8**
Space and Furnishings Subscale	**5.0**	**0.9**	**3.0-6.8**	**3.9**	**0.7**	**2.6-5.8**
Indoor space	5.0	1.9	1-7	4.6	1.6	2-7
Furniture for routine care, play, and learning	6.4	1.2	2-7	5.1	1.5	2-7
Furnishings for relaxation and comfort	5.5	1.6	3-7	5.0	1.8	1-7
Room arrangement for play	5.6	1.7	1-7	3.3	1.7	2-7
Space for privacy	5.2	1.9	2-7	3.5	1.9	2-7
Child-related display	4.9	1.5	3-7	4.6	1.5	2-7
Space for gross motor play	3.5	2.0	1-7	1.8	1.3	1-7
Gross motor equipment	3.9	2.3	1-7	3.2	2.0	1-7
Personal Care Routines Subscale	**4.9**	**1.1**	**2.3-7.0**	**2.8**	**0.9**	**1.3-5.7**
Greeting/departing	6.6	0.9	4-7	5.5	1.9	1-7
Meals/snacks	4.0	2.1	1-7	1.8	1.1	1-6
Nap/rest[c]	5.0	2.0	2-7	2.8	2.0	1-7
Toileting/diapering	5.1	2.5	1-7	2.4	1.6	1-7
Health practices	5.2	1.9	1-7	2.7	1.7	1-7
Safety practices	3.9	2.5	1-7	1.4	0.6	1-4

[a] The Total Score includes all items on the ECERS-R (items 1-43).
[b] The Total Child Items Score includes all items on the ECERS-R but the Parents and Staff subscale (items 1-37).
[c] For this item in 2005-2006, n=56.

Table 10. Quality of Classroom Practices (ECERS-R) (continued)

Item Description	2003-2004 (n=99)			2005-2006 (n=57)		
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**
Language-Reasoning Subscale	**5.8**	**0.9**	**3.3-7.0**	**4.8**	**0.8**	**3.3-7.0**
Books and pictures	5.5	1.5	2-7	4.3	1.3	1-7
Encouraging children to communicate	6.6	0.8	4-7	6.3	1.0	4-7
Using language to develop reasoning skills	4.9	1.5	2-7	4.1	1.2	2-7
Informal use of language	5.9	1.4	2-7	4.4	1.1	3-7
Activities Subscale	**4.9**	**0.8**	**2.8-6.6**	**4.5**	**0.9**	**2.2-6.9**
Fine motor	5.6	1.5	3-7	5.2	1.4	2-7
Art	5.0	1.7	1-7	4.4	1.5	2-7
Music/ movement	4.3	1.6	2-7	4.7	1.5	2-7
Blocks	4.5	1.1	3-7	4.3	1.2	1-7
Sand/water	4.8	1.4	1-7	5.4	1.6	1-7
Dramatic play	4.9	1.4	2-7	4.6	1.1	2-7
Nature/science	4.5	1.7	2-7	4.3	1.4	2-7
Math/number	4.9	1.5	1-7	4.5	1.4	1-7
Use of TV, video, and/or computers[a]	5.2	2.0	1-7	3.7	2.0	1-7
Promoting acceptance of diversity	5.1	1.4	2-7	4.2	1.8	2-7

[a] For this item in 2003-2004, n=90 and in 2005-2006, n=55.

Table 10. Quality of Classroom Practices (ECERS-R) (continued)

Item Description	2003-2004 (n=99)			2005-2006 (n=57)		
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**
Interaction Subscale	**6.2**	**1.0**	**1.4-7.0**	**4.8**	**1.2**	**2.0-7.0**
Supervision of gross motor activities[a]	5.1	1.7	1-7	4.2	1.4	2-7
General supervision of children	6.3	1.4	1-7	4.6	2.0	1-7
Discipline	6.2	1.2	1-7	4.6	1.6	1-7
Staff-child interactions	6.6	1.2	1-7	5.3	2.0	1-7
Interactions among children	6.6	1.0	1-7	5.4	1.7	2-7
Program Structure Subscale	**6.2**	**0.9**	**3.8-7.0**	**4.4**	**1.4**	**1.7-7.0**
Schedule	6.0	1.6	2-7	2.9	1.5	2-7
Free play	6.3	1.3	1-7	4.8	2.3	2-7
Group time	6.3	1.2	3-7	4.9	1.9	1-7
Provisions for children with disabilities[b]	6.1	1.2	1-7	5.8	1.5	2-7
Parents and Staff Subscale	**5.3**	**0.9**	**2.5-7.0**	**5.6**	**0.8**	**3.0-7.0**
Parent provisions	5.9	1.1	1-7	5.9	1.1	3-7
Staff personal needs, provisions	3.4	1.6	1-7	3.4	1.7	1-7
Staff professional needs, provisions	4.8	2.1	1-7	5.3	2.3	1-7
Staff interaction	6.6	1.1	1-7	6.4	1.1	2-7
Supervision/evaluation of staff	5.9	1.5	1-7	6.3	1.3	2-7
Professional growth opportunities	5.3	1.6	1-7	6.1	1.3	2-7

[a] For this item in 2003-2004, n=98.
[b] For this item in 2003-2004, n=70 and in 2005-2006, n=40.

Figure 2. Distribution of Classroom Practices Scores (ECERS-R Total Mean Item Scores)

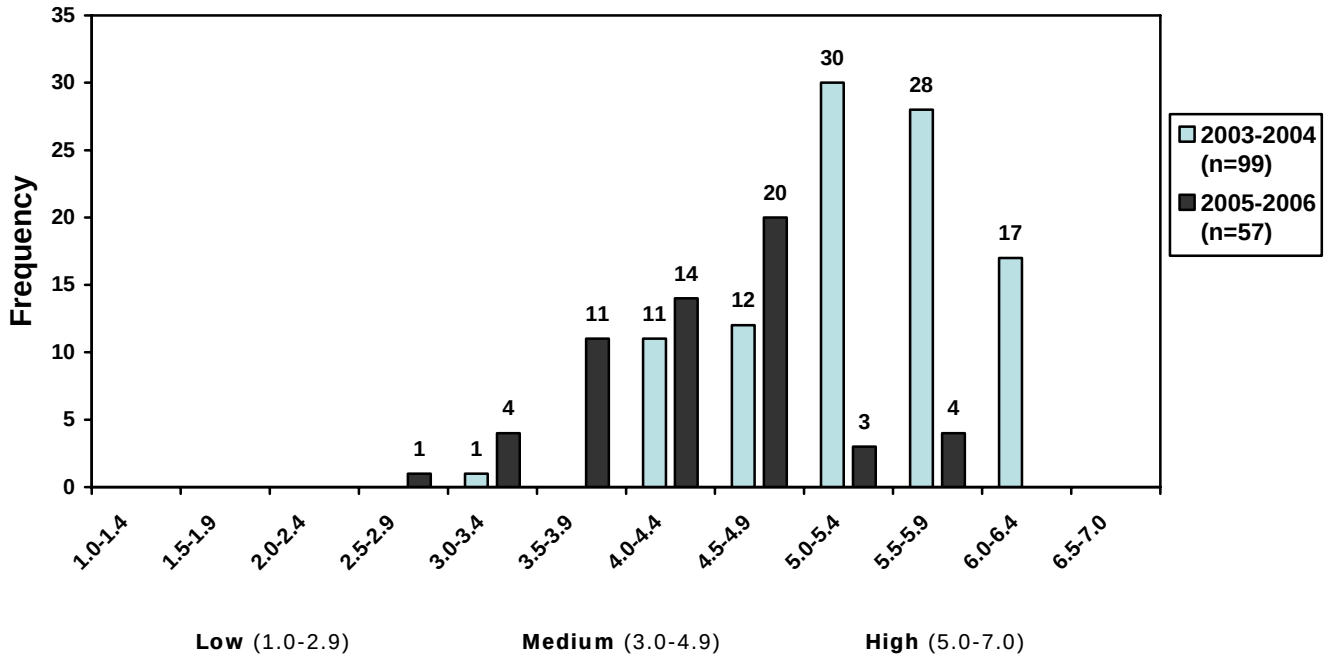
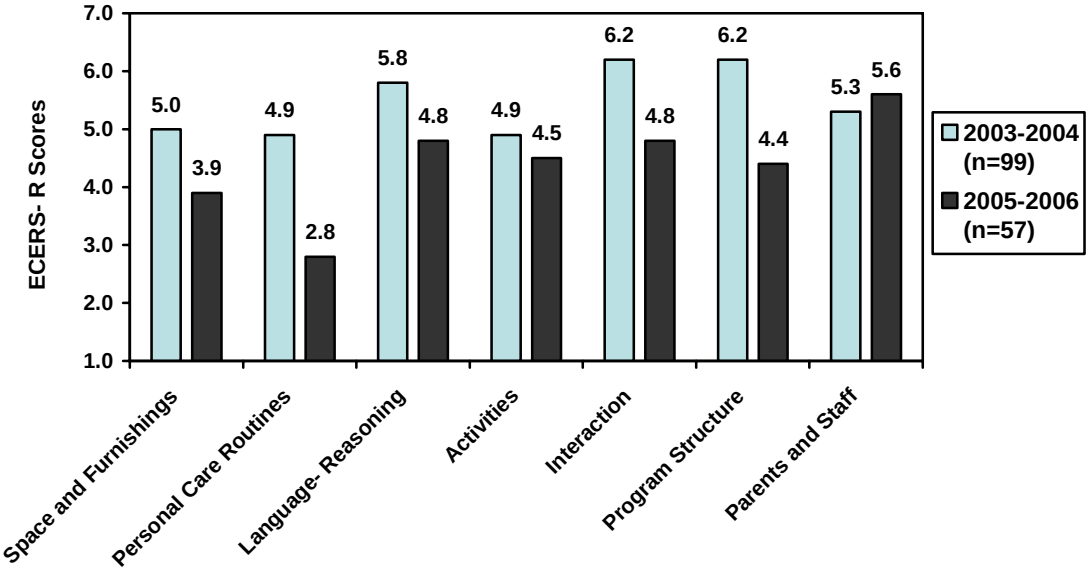


Figure 3. Classroom Practices Mean Subscale Scores (ECERS-R)



Literacy Environment

Observations of the quality of the literacy environment were conducted for the 2005-2006 cohort using the ELLCO (see Table 11, Figure 4, Figure 5, and Figure 6). The Classroom Observation Scale, which includes subscales measuring both the General classroom environment and the Language/literacy/curriculum environment, is the primary quality indicator on the ELLCO. Items on this scale are scored from 1-5, representing quality levels from deficient (1) to basic (3) to exemplary (5). The average item score on the Classroom Observation Scale was 3.7, approaching the exemplary side of the scale. Scores were somewhat higher on the General classroom environment subscale (4.0) than on the Language, literacy and curriculum subscale (3.6).

Comparisons across the different scales on the ELLCO indicate that, in general, scores were relatively higher on the Classroom Observation Scale and the Literacy Environment Checklist (which measures the availability of books and writing materials in the room) than on the Literacy Activities Rating Scale (which measures the frequency of book reading and writing activities). Average scores on the Classroom Observation Scale represented 74% of the total possible and 71% on the Literacy Environment Checklist, compared to 63% for the Literacy Activities Rating Scale. These findings suggest that the More at Four classrooms are doing a somewhat better job of setting up a literacy-rich environment than actually carrying out literacy-related activities, with writing activities scoring proportionally lower than reading activities.

Table 11. Quality of Literacy Environment (ELLCO)

(2005-2006 sample)
n=57

Item Description	Mean	SD	Range	Total Possible Range
Classroom Observation Scale (Mean Item Score)	**3.7**	**0.6**	**2.5-4.9**	**1-5**
General Classroom Environment	4.0	0.7	2.4-5.0	1-5
Language, Literacy and Curriculum	3.6	0.7	2.0-4.8	1-5
Literacy Environment Checklist (Total Score)	**29.2**	**5.8**	**17-40**	**0-41**
Book Area	2.3	0.7	0-3	0-3
Book Selection	7.5	0.7	5-8	0-8
Book Use	5.2	2.7	0-9	0-9
Writing Materials	6.2	1.1	4-8	0-8
Writing Around the Room	8.0	2.9	1-13	0-13
Literacy Activities Rating Scale (Total Score)	**8.2**	**2.3**	**3-12**	**0-13**
Book Reading	5.3	1.8	2-8	0-8
Writing	2.9	1.6	0-5	0-5

Figure 4. Distribution of Classroom Observation Scale Scores (ELLCO)
(2005-2006 Sample)

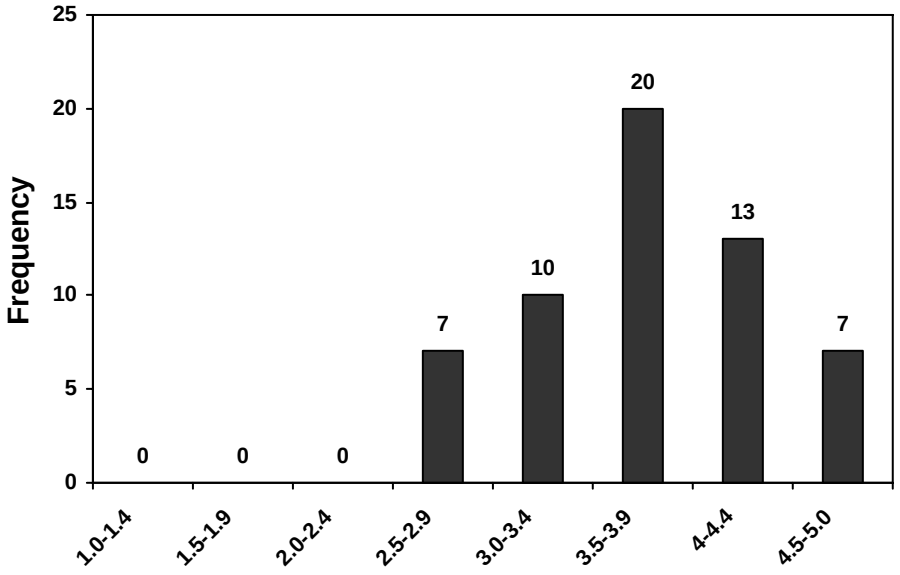


Figure 5. Distribution of Literacy Environment Checklist Scores (ELLCO)
(2005-2006 Sample)

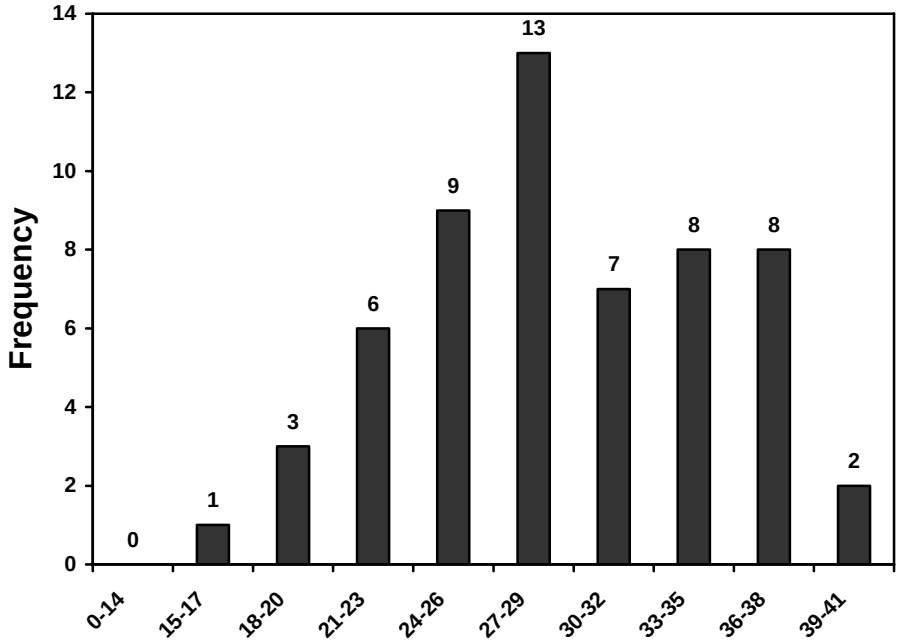
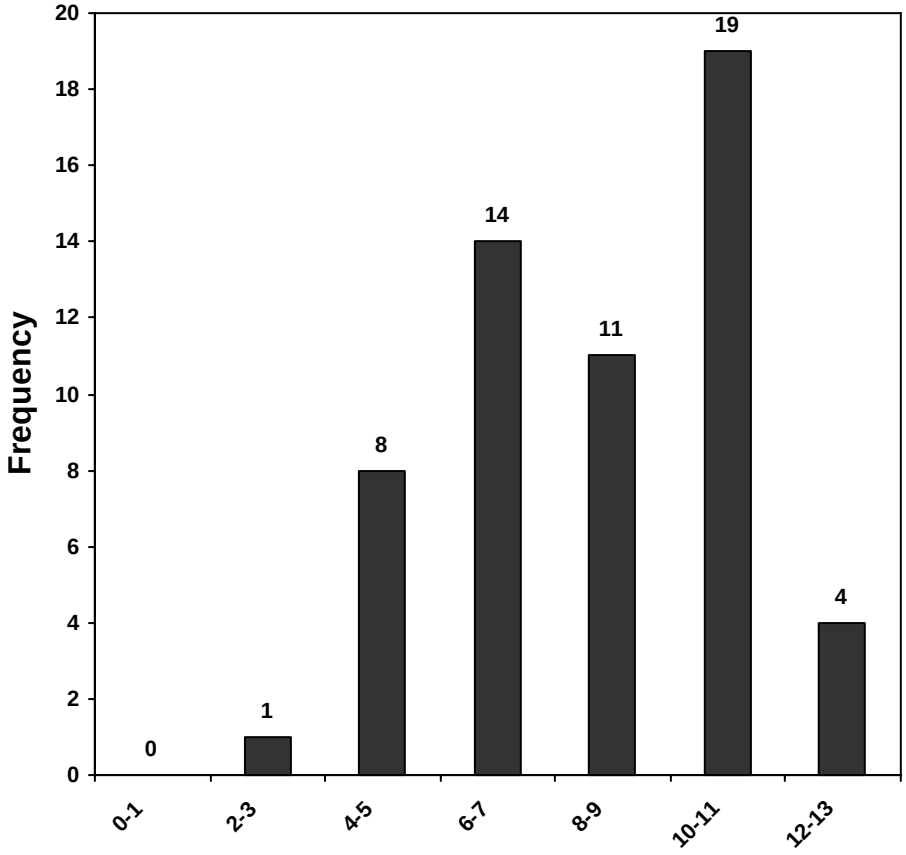


Figure 6. Distribution of Literacy Activities Rating Scale Scores (ELLCO)
(2005-2006 Sample)



Teacher-Child Interactions

Observations of the quality of teacher-child interactions were conducted for the 2005-2006 cohort using the CIS (see Table 12, Figure 7, and Figure 8). Average scores on the CIS indicate that teachers were fairly sensitive in their interactions with children. The majority of classrooms (88%) scored 3.0 or above on the CIS total score, with higher scores representing more positive interactions. Scores on the Sensitivity subscale, which indicates positive interactions with children, were fairly high, while scores on the Harshness, Detachment, and Permissiveness subscales, which indicate negative interactions, were fairly low (i.e., fewer negative interactions occurred). On the Sensitivity subscale, 58% of the classrooms scored 3.0 or above, where higher scores represent more positive interactions. For the three negative subscales, where lower scores represent more positive interactions, 84% scored below 2.0 and 23% had scores of 1.0 on Harshness, 96% scored below 2.0 and 63% had scores of 1.0 on Detachment, and 88% scored below 2.0 and 35% had scores of 1.0 on Permissiveness.

**Table 12. Quality of Teacher-Child Interactions (Caregiver Interaction Scale)
(2005-2006 Sample)
n=57**

Item Description[a]	Mean	SD	Range[b]
Total Items Score	**3.4**	**0.4**	**2.4-3.9**
Sensitivity Subscale	3.1	0.4	2.2-3.8
Harshness Subscale	1.5	0.5	1.0-3.3
Detachment Subscale	1.2	0.3	1.0-2.3
Permissiveness Subscale	1.4	0.4	1.0-2.3

[a] For the total score calculation, scoring is reversed on the Harshness, Detachment, and Permissiveness subscales so that higher total scores represent more positive interactions. For the individual scores on these three subscales, lower scores represent more positive interactions, while for the Sensitivity subscale, higher scores represent more positive interactions.
[b] Possible range=1-4.

Figure 7. Distribution of Teacher-Child Interaction Scores (CIS Total)
(2005-2006 Sample)

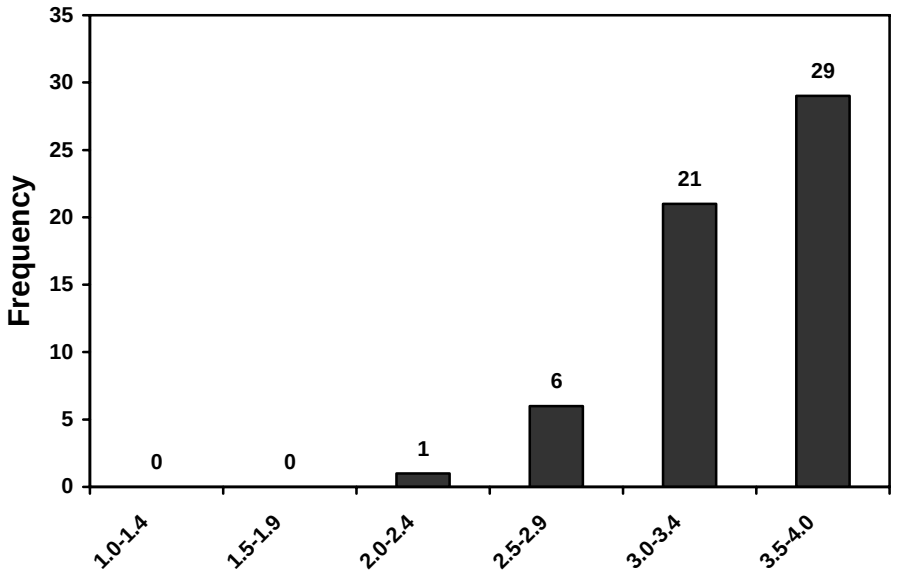
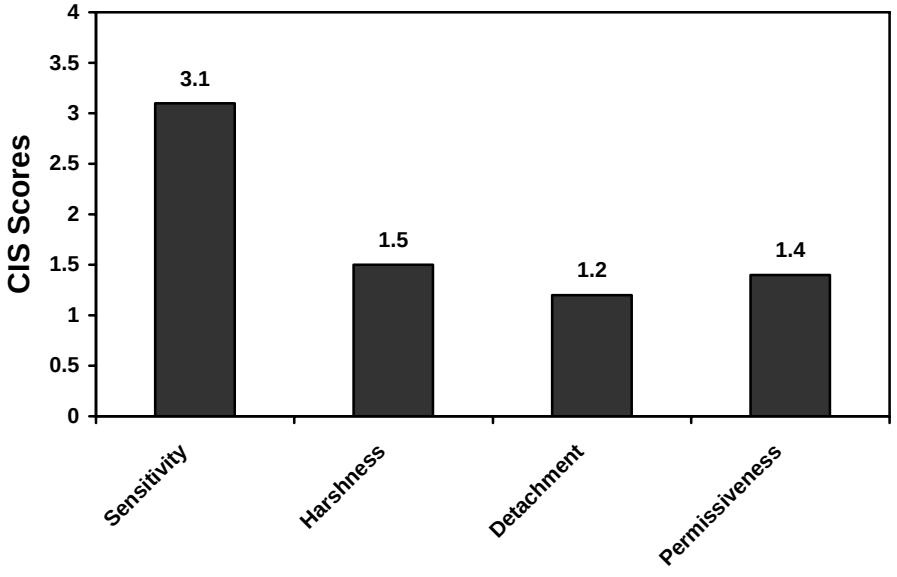


Figure 8. Teacher-Child Interaction Mean Subscale Scores (CIS)[a]
(2005-2006 Sample)



[a] Note: Higher scores on the Sensitivity subscale and lower scores on the Harshness, Detachment, and Permissiveness subscales represent higher quality interactions.

Factors Predicting Classroom Quality

Three sets of factors were examined hierarchically to determine whether they were associated with differences in classroom quality: 1) Teacher qualifications (whether or not the lead teacher had a B-K license or the equivalent); 2) Class size (averaged across each month of program operations); and 3) Classroom-level characteristics of the children served, including average proportion of More at Four children, average risk total score for More at Four children, and average service priority status for More at Four children. Each of the three classroom quality measures was examined separately.

Classroom Practices

For the 2005-2006 cohort, higher classroom quality scores on the ECERS-R (total child items score) were associated with having a lower proportion of More at Four children in the classroom [ß(se)=-0.92 (.43), $p<.04$]. None of the other factors were related to the quality of classroom practices.

For classroom practices, a second set of analyses was performed to examine whether these findings were consistent across the 2003-2004 and 2005-2006 cohorts. These analyses excluded average service priority status because of changes in the definitions in the program guidelines over this period. When the classroom quality data were combined across the two cohorts, none of the factors examined were associated with classroom quality. However, there was a significant effect of cohort, indicating that scores were significantly higher in 2003-2004 than in 2005-2006 [$F(1,150)=103.44$, $p<.0001$].

Literacy Environment

Analyses of the 2005-2006 data were conducted to examine whether the quality of the literacy environment (ELLCO Classroom Observation score) was associated with teacher qualifications, class size, or classroom-level characteristics of the children served. These results indicated that higher quality literacy environments were associated with lead teachers having a B-K license (or equivalent) even after adjusting for class size, but this effect was not maintained after adjusting for other classroom characteristics (average proportion of More at Four children, average risk total score, and average service priority status).

Sensitivity of Teacher-Child Interactions

The 2005-2006 data were also analyzed to examine whether the sensitivity of teacher-child interactions (CIS total score) was associated with teacher qualifications, class size, or classroom-level characteristics of the children served. These analyses revealed that there were no significant associations between any of these teacher or classroom factors and the quality of teacher-child interactions.

Analysis Strategies

To examine the factors predicting classroom quality, we conducted a series of hierarchical regression models using a general linear models approach. Three blocks of variables were added hierarchically to the model: 1) Teacher qualifications, a 2-level categorical variable measuring whether or not the teacher had a B-K license or the equivalent; 2) Average class size, measured as the average of the monthly reported class size; and 3) Classroom-level child characteristics, including average proportion of More at Four children (measured as the average monthly reported proportion of More at Four children in each classroom), average risk total (measured as the average of the reported risk total score at entry into the program for all More at Four children in the class), and average service priority status (measured as the average of the reported service priority level at entry into the program for all More at Four children in the class).

For the analyses of the combined cohorts, average service priority status was excluded because of changes in the definitions in the program guidelines between the two cohorts. In addition, cohort was included as an independent variable in these analyses. Further, the interactions between cohort and the other predictors in the model were tested to examine whether there were any differences in these factors by cohort; however, none of these interactions were significant, so they were dropped from the final models.

Separate analyses were conducted for each aspect of classroom quality: 1) Classroom practices, as measured by the total child items score on the ECERS-R; 2) Literacy environment, as measured by the Classroom Observation Scale score on the ELLCO; and 3) Sensitivity of teacher-child interactions, as measured by the total score on the CIS. For the combined cohorts, analyses were only conducted for classroom practices, since that was the only quality measure available across both years.

CHILD OUTCOMES

In order to address questions about the outcomes for children attending More at Four and factors associated with better outcomes, individual child assessments were conducted near the beginning and end of the program year in 2003-2004 and 2005-2006. The child assessments included measures of children's language and literacy skills (receptive language, rhyming, story concepts, letter naming), math skills (applied problems, counting), general knowledge (color knowledge, social awareness), and behavioral skills (social skills, problem behaviors). Two sources of data were gathered: trained assessors administered measures of children's language/literacy skills, math skills, and general knowledge, and teachers completed ratings of children's behavioral skills. In addition, for the 2005-2006 sample only, assessments were administered in both English and Spanish for Spanish-speaking children.

These data provided information about the amount of developmental growth experienced by children over the More at Four program year based on a number of widely-used measures. In accord with the overall goal of More at Four, the outcome areas measured were consistent with generally accepted definitions of school readiness, including the recommendations of the National Education Goals Panel.[7]

We conducted six series of analyses to examine children's developmental growth over the More at Four year and factors affecting their outcomes. For each series of analyses, we conducted single cohort analyses for the 2005-2006 sample and compared these findings to analyses of the 2003-2004 cohort and/or the combined data from the two cohorts. The first series of analyses examined whether there were significant increases over time in children's skill levels. The second series of analyses examined whether classroom quality characteristics (classroom practices, literacy environment, teacher-child interactions) were associated with differences in children's outcomes. The third, fourth, and fifth series of analyses examined whether differences in the level of individual child risk characteristics (cumulative risk level and English proficiency level) were associated with differences in children's development. In addition, a sixth set of analyses was conducted to examine growth on the same child outcome measures in Spanish as well as English for a subset of Spanish-speaking children in the 2005-2006 cohort.

Changes over Time in Child Outcomes

The first series of analyses examined the amount of growth children exhibited on the various outcome measures over the More at Four program year. For these analyses, we tested whether the amount of growth using change scores (spring score minus fall score) was significantly different from zero, adjusting for classroom to account for multiple children in each classroom.

Single Cohort Analyses

As indicated in Table 13, the change scores were significantly different from zero for all of the language and literacy measures (receptive language, rhyming, story and print concepts, naming letters); the math measures (applied problems, counting task); the general knowledge measures (color knowledge, social awareness); and the social skills measure for the 2005-2006 cohort as well as the 2003-2004 cohort. These findings indicate that children made significant gains in all skill areas over the course of the More at Four year, and are consistent with the findings from

previous years. The one area that showed no significant change was children's problem behaviors, which were slightly below the population mean (slightly fewer problem behaviors), which is again consistent with the findings from previous years.

Combined Cohort Analyses

The data from the 2003-2004 and 2005-2006 pre-k cohorts were combined to determine whether the findings were consistent when multiple years of the More at Four Program were examined. A similar series of analyses was conducted as described previously for the single cohort analyses to examine whether there were significant increases over time in children's skill levels, as well as whether there were any differences by cohort (2003-04 vs. 2005-06). For these analyses, we tested whether the change scores (spring score minus fall score) were significantly different from zero. As seen in Table 13, for the combined cohorts, children made significant gains during the program year on all outcome measures except for problem behaviors, which remained just below the expected population mean (i.e., slightly fewer problem behaviors than expected). These results mirror the results for the individual cohorts, indicating that these findings are robust over time.

Table 13. Child Outcome Scores by Year

Domain	Outcome	2003-2004		2005-2006		Significance of Change for Combined Cohort[a,b]
		Fall (n=453-514) **Mean** (SD) Range	**Spring**[a,b] (n=419-464) **Mean** (SD) Range	**Fall** (n=416-478) **Mean** (SD) Range	**Spring**[a,b] (n=372-445) **Mean** (SD) Range	
Language and literacy	Receptive Language (PPVT-III[c])	**85.4** (19.3) 40-124	**89.9**** (17.2) 40-126	**81.1** (20.9) 21-125	**87.0***** (19.6) 32-129	***
	Rhyming (WJ-III[d])	**1.9** (2.7) 0-15	**4.4***** (4.1) 0-15	**1.9** (2.8) 0-15	**3.8***** (3.8) 0-15	***
	Story and Print Concepts[e]	**3.0** (2.2) 0-10	**4.9***** (2.6) 0-12	**2.9** (2.4) 0-10	**4.7***** (2.6) 0-12	***
	Naming Letters[f]	**6.1** (7.9) 0-26	**15.1***** (9.5) 0-26	**6.6** (8.6) 0-26	**15.3***** (9.6) 0-26	***
Math	Applied Problems (WJ-III[c])	**93.1** (15.0) 46-128	**94.0**** (13.9) 51-124	**91.1** (15.9) 47-135	**94.0***** (14.3) 46-127	***
	Counting Task[g]	**11.8** (8.1) 1-40	**18.9***** (11.5) 1-40	**11.5** (7.9) 1-40	**18.9***** (10.6) 2-40	***
General knowledge	Social Awareness[h]	**3.7** (1.8) 0-6	**4.5***** (1.5) 0-6	**3.3** (1.9) 0-6	**4.2***** (1.5) 0-6	***
	Color Knowledge[i]	**16.3** (5.6) 0-20	**18.8***** (2.7) 3-20	**15.6** (6.0) 0-20	**18.6***** (3.2) 0-20	***
Classroom behavior	Social Skills (SSRS[c])	**100.8** (15.3) 54-130	**107.8***** (15.3) 60-130	**100.4** (15.7) 53-130	**109.7***** (14.7) 60-130	***
	Problem Behaviors (SSRS[c])	**98.6** (11.9) 85-138	**99.3**[NS] (12.8) 85-145	**98.2** (13.1) 85-142	**97.2**[NS] (12.0) 85-135	NS

[a] *p < .05, **p < .01, ***p <.001, NS=nonsignificant.
[b] Significance levels indicate results of test of whether change scores are different from zero, adjusting for classroom.
[c] Indicates standardized, norm-referenced measure with mean=100, SD=15.
[d] Possible range=0-17.
[e] Possible range=0-14.
[f] Possible range=0-26.
[g] Possible range=0-40.
[h] Possible range=0-6.
[i] Possible range=0-20.

Differences in Child Outcomes by Classroom Quality

For the second series of analyses, we examined whether different aspects of classroom quality were associated with differences in children's rate of growth over the More at Four program year, using change scores on the individual child outcome measures. Separate analyses were conducted for the three measures of classroom quality: 1) overall classroom practices as measured by the ECERS-R total child items score, 2) literacy environment as measured by the Classroom Observation Scale of the ELLCO, and 3) the sensitivity of teacher-child interactions as measured by the CIS total score. For these analyses, we adjusted for children's risk factor level, assessed English proficiency level, age at fall assessment, gender, and attendance.

2005-2006 Cohort. In general, there were few effects related to differences in classroom quality for the 2005-2006 sample. For the classroom practices measure, better quality was associated with lower gains on letter naming [ß (se)=-1.53 (.76), p<.05]. For the literacy environment measure, higher quality was associated with greater gains on the rhyming task [ß (se)=.90 (.36), p<.05]. For the teacher-child interaction measure, higher quality was associated with greater gains in social skills [ß (se)=6.31 (3.07), p<.05].

Combined Cohort. These analyses examined the associations between the quality of classroom practices (as measured by the ECERS-R) and children's rate of growth for the combined cohorts. (The other two measures of classroom quality were not gathered for the 2003-2004 sample.) There were no significant associations between the quality of classroom practices and children's rate of growth over the program year when the two cohorts were combined. These findings contrast with the single cohort analyses, which indicated one negative association between the quality of classroom practices and children's gains (letter naming), as well as positive associations for the quality of the literacy environment (rhyming) and teacher-child interactions (social skills).

Differences in Child Outcomes by Cumulative Risk Factor Levels

For the third series of analyses, we examined whether there were significant differences in the level of skills or the rate of growth for children entering the program at different levels of cumulative risk. Children were categorized according to four levels of cumulative risk (0-3 from low risk to high risk) based on poverty level (eligibility for free lunch, reduced-price lunch, or full-price lunch) and presence or absence of an identified special need, limited English proficiency, and chronic health condition.[a] To examine differences in the absolute level of scores by risk factor groups, separate analyses were conducted for the fall and spring scores. To examine differences in the amount of growth for children in different risk factor groups, analyses of the change scores (spring scores minus fall scores) were conducted. In addition, these analyses adjusted for children's gender and attendance.

[a] A total risk factor score was constructed based on More at Four eligibility guidelines, using income (eligibility for free lunch=2 points, reduced-price lunch=1 point, and full-price lunch=0 points) and additional risk factors (1 point each for limited English proficiency, identified disability, and chronic health condition). A four-level categorical variable was constructed, representing risk factor scores of 0, 1, 2, and 3-5. Presence or absence of developmental/educational need (included in the 2005-2006 program guidelines) was not included in the calculation of risk factor total in order to be consistent with previous years.

Differences in Level of Skills by Risk Factor Group

As seen in Table 14, children in the highest-risk group were scoring significantly lower than children in the other risk groups in both the fall and spring on nearly all the language/literacy (receptive language, rhyming, story concepts, letter naming), math skills (applied problems, counting), and general knowledge (social awareness) measures for the 2005-2006 cohort. On one measure, color knowledge, these differences were not significant by the spring for this sample, indicating that children in the highest-risk group had caught up to the other groups by the end of the pre-k year. Children in all groups were scoring at or near the ceiling (maximum score) on this measure by the spring. There were no differences among risk groups for measures of behavioral skills (social skills, problem behaviors). A similar pattern of results is seen for the 2003-2004 cohort, although the differences for color knowledge remained significant in the spring.[a]

Differences in Rate of Growth by Risk Factor Group

2005-2006 Cohort. These analyses examined whether there were differences in the rate of growth experienced by children at different levels of risk for the 2005-2006 sample. Children in the highest-risk group made significantly greater gains on applied math skills [$F(3, 74)=2.75$, $p<.05$] and color knowledge [$F(3, 86)=7.77$, $p<.001$] than children in the other three groups. For color knowledge, all groups were scoring near the ceiling by the spring, although children at greatest risk started lower in the fall. These differences are illustrated in Figure 9 and Figure 10, which compare performance for the highest- and lowest-risk groups. In contrast, children in the lowest-risk group made significantly greater gains than other children on rhyming skills [$F(3, 85)=3.44$, $p<.05$], an area of phonological awareness that requires higher-level abilities (see Figure 11). There were no differences in the amount of gain on other measures of language/literacy skills (receptive language, story concepts, letter naming), math skills (counting), general knowledge (social awareness), or behavioral skills (social skills, problem behaviors).

[a] For both cohorts, we also examined whether children at greater risk were in classrooms with lower average quality or higher average proportions of More at Four children, to determine whether there were differences in the quality of their experiences. We could not statistically test these differences because there were children at different risk levels within a classroom, and therefore, the measurement of risk level categories and classroom characteristics were not independent. However, examination of the means for each risk group provided no evidence that the most at-risk children were in lower-quality classrooms for either cohort; if anything, slightly the reverse was true (2003-2004 range=5.1-5.5 and 2005-2006 range=3.9-4.2). Similarly, there were only slight differences in proportion of More at Four children by risk level groups, with no evidence that children in the highest risk group were in classrooms with the highest proportions (2003-2004 range=.84-.89 and 2005-2006 range=.84-.92).

Combined Cohort. The purpose of these analyses was to examine whether risk factor group was a significant predictor of children's rate of growth for the combined sample, and whether these findings were consistent with the single cohort results. For the combined sample, children at greater risk showed greater gains in receptive language [$F(3, 162)=4.57$, $p<.01$], applied math skills [$F(3, 143)=6.85$, $p<.001$], color knowledge [$F(3, 165)=25.38$, $p<.001$], and social awareness [$F(3, 164)=6.48$, $p<.001$]. In contrast, lower-risk children showed greater gains on rhyming [$F(3,157)=4.12$, $p<.01$]. There were no differences on the basis of risk for other language/literacy skills (story concepts and letter naming), math skills (counting), or behavioral skills (social skills and problem behaviors). This pattern of results is largely consistent with the findings from the 2005-2006 single cohort analyses reported previously. The two exceptions are receptive language and social awareness, both of which were not significant for the single cohort analyses, but reached significance in the combined sample. In addition, these analyses revealed some differences in the overall level of gain by cohort. The 2005-2006 cohort showed significantly greater gains than the 2003-2004 cohort for receptive language [ß (se)=2.56 (.84), $p<.01$] and applied math skills [ß (se)=2.39 (1.06), $p<.05$]. There were no cohort differences for the remaining measures.

Table 14. Child Outcomes by Risk Factor Levels

Outcome	Risk Total Group[a]	2003-2004			2005-2006		
		N[b]	Fall	Spring	N[c]	Fall	Spring
			Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)
Language and Literacy							
Receptive Language (PPVT-III[d])	**0**	35	98.2 (15.6)	97.8 (12.2)	27	94.1 (10.8)	99.6 (10.9)
	1	56	92.9 (12.4)	96.6 (12.0)	66	87.5 (16.9)	93.4 (17.6)
	2	258	89.1 (15.1)	93.0 (14.0)	238	85.1 (19.4)	91.0 (17.3)
	3+	78	65.2 (21.1)	72.6 (19.3)	95	65.2 (19.8)	73.9 (18.5)
	Significant group differences[e]:		3<0,1,2 2<0	3<0,1,2	Sig[e]:	3<0,1,2	3<0,1,2
Rhyming (WJ-III[f])	**0**	33	4.0 (4.1)	6.8 (4.9)	26	2.7 (2.9)	6.2 (4.0)
	1	51	2.1 (2.4)	4.6 (4.1)	68	2.7 (3.6)	4.8 (4.3)
	2	251	2.1 (2.8)	4.9 (4.2)	241	2.1 (3.0)	4.1 (3.9)
	3+	56	0.6 (1.2)	2.2 (2.3)	104	0.6 (1.3)	2.1 (2.4)
	Significant group differences[e]:		3<0,1,2 2<0	3<0,1,2	Sig[e]:	3<0,1,2	3<0,1,2

[a] A total risk factor score was constructed based on More at Four eligibility guidelines, using income (eligibility for free lunch=2 points, reduced-price lunch=1 point, and full-price lunch=0 points) and additional risk factors (1 point each for limited English proficiency, identified disability, and chronic health condition). A four-level categorical variable was constructed, representing risk factor scores of 0, 1, 2, and 3-5.
[b] A total of 514 children were assessed in 2003-2004, but analyses for each measure were only performed for those with both fall and spring data.
[c] A total of 478 children were assessed in 2005-2006, but analyses for each measure were only performed for those with both fall and spring data.
[d] Indicates standardized, norm-referenced measure with mean=100, SD=15.
[e] Significant comparisons reported represent significant group differences based on all possible pairwise comparisons of risk factor groups using adjusted p-values for multiple comparisons.
[f] Possible range=0-17.

Table 14. Child Outcomes by Risk Factor Levels (continued)

Outcome	Risk Total Group	2003-2004			2005-2006		
		N	Fall	Spring	N	Fall	Spring
			Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)
Language and Literacy							
Story and Print Concepts[a]	**0**	35	4.5 (2.5)	6.2 (2.5)	26	3.7 (2.6)	5.5 (2.3)
	1	54	3.5 (2.3)	5.5 (2.8)	68	3.6 (2.5)	5.6 (2.6)
	2	256	3.1 (2.1)	5.0 (2.4)	241	3.1 (2.4)	4.8 (2.6)
	3+	64	2.1 (1.9)	4.2 (2.6)	103	1.9 (1.8)	3.6 (2.3)
	Significant group differences:		3<0,1,2 2<0	3<0,1	Sig:	3<0,1,2	3<0,1,2
Naming Letters[b]	**0**	35	10.4 (9.8)	17.7 (9.1)	27	9.2 (8.5)	17.3 (8.5)
	1	54	7.3 (9.1)	15.5 (9.0)	68	9.1 (9.8)	17.4 (9.1)
	2	259	6.7 (8.0)	15.9 (9.5)	246	6.8 (8.7)	16.0 (9.4)
	3+	79	3.2 (5.8)	11.0 (9.1)	104	3.4 (6.4)	11.7 (10.0)
	Significant group differences:		3<0,1,2	3<0,1,2	Sig:	3<1	3<0,1,2

[a] Possible range=0-14.
[b] Possible range=0-26.

Table 14. Child Outcomes by Risk Factor Levels (continued)

Outcome	Risk Total Group	2003-2004			2005-2006		
		N	Fall	Spring	N	Fall	Spring
			Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)
Math							
Applied Problems (WJ-III[a])	**0**	33	99.6 (15.2)	100.2 (10.8)	26	99.8 (12.6)	102.1 (10.3)
	1	52	99.3 (13.0)	98.8 (11.9)	61	97.4 (13.6)	100.8 (10.4)
	2	245	93.7 (13.6)	95.9 (12.7)	219	92.3 (14.5)	96.1 (12.1)
	3+	51	81.6 (16.4)	88.0 (13.4)	76	81.6 (16.2)	91.0 (14.9)
	Significant group differences:		3<0,1,2 2<1	3<0,1,2	Sig:	3<0,1,2	3<0,1,2
Counting Task[b]	**0**	33	16.9 (9.9)	22.5 (12.0)	26	13.4 (8.2)	25.0 (11.1)
	1	53	12.4 (6.1)	20.0 (12.3)	67	14.2 (8.8)	22.1 (11.1)
	2	250	12.7 (8.9)	20.4 (11.6)	238	11.7 (7.9)	19.0 (10.5)
	3+	68	8.0 (4.7)	14.6 (8.5)	98	9.0 (6.1)	16.4 (9.3)
	Significant group differences:		3<0,1,2	3<0,1,2	Sig:	3<1	3<0,1 2<0

[a] Indicates standardized, norm-referenced measure with mean=100, SD=15.
[b] Possible range=0-40.

Table 14. Child Outcomes by Risk Factor Levels (continued)

Outcome	Risk Total Group	2003-2004			2005-2006		
		N	Fall	Spring	N	Fall	Spring
			Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)
General Knowledge							
Color Knowledge[a]	**0**	36	18.5 (2.5)	19.7 (0.6)	27	18.6 (3.1)	18.8 (1.9)
	1	56	17.7 (4.3)	19.3 (2.0)	68	16.5 (5.8)	19.0 (1.8)
	2	261	17.1 (5.0)	18.9 (2.5)	246	16.3 (5.5)	18.7 (3.2)
	3+	79	12.0 (6.8)	17.6 (3.7)	104	13.2 (6.5)	18.1 (4.1)
	Significant group differences:		3<0,1,2	3<0,1,2	Sig:	3<0,1,2	NS
Social Awareness[b]	**0**	36	4.6 (1.6)	5.3 (1.1)	27	4.0 (1.4)	4.8 (0.9)
	1	56	4.0 (1.3)	5.0 (1.2)	67	3.8 (1.8)	4.6 (1.3)
	2	261	4.0 (1.6)	4.6 (1.4)	244	3.7 (1.8)	4.4 (1.5)
	3+	80	2.0 (1.5)	3.3 (1.6)	103	2.0 (1.6)	3.2 (1.5)
	Significant group differences:		3<0,1,2	3<0,1,2	Sig:	3<0,1,2	3<0,1,2

[a] Possible range=0-20.
[b] Possible range=0-6.

Table 14. Child Outcomes by Risk Factor Levels (continued)

Outcome	Risk Total Group	2003-2004			2005-2006		
		N	Fall	Spring	N	Fall	Spring
			Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)
Behavioral Skills							
Social Skills (SSRS[a])	**0**	37	105.5 (13.8)	109.2 (15.9)	18	107.5 (10.3)	114.4 (11.6)
	1	61	101.2 (17.3)	105.8 (15.1)	60	101.8 (13.9)	109.6 (13.4)
	2	274	101.1 (14.8)	108.2 (14.2)	204	100.9 (15.4)	108.6 (15.6)
	3+	87	99.8 (16.3)	107.7 (18.4)	79	102.7 (15.3)	113.3 (13.2)
	Significant group differences:		3<2	NS	Sig:	NS	NS
Problem Behaviors (SSRS[a])	**0**	38	100.4 (12.4)	98.7 (13.6)	20	94.2 (9.7)	95.0 (9.4)
	1	61	98.6 (12.3)	101.6 (13.0)	62	95.9 (12.1)	98.0 (12.4)
	2	274	98.0 (12.1)	99.0 (12.6)	208	98.8 (13.6)	98.1 (12.6)
	3+	87	98.4 (11.4)	98.2 (12.8)	83	96.0 (12.5)	94.8 (10.6)
	Significant group differences:		NS	NS	Sig:	NS	NS

[a] Indicates standardized, norm-referenced measure with mean=100, SD=15.

Figure 9. Growth in Math Skills (WJ-III Applied Problems) by Cumulative Risk (2005-2006 Cohort)

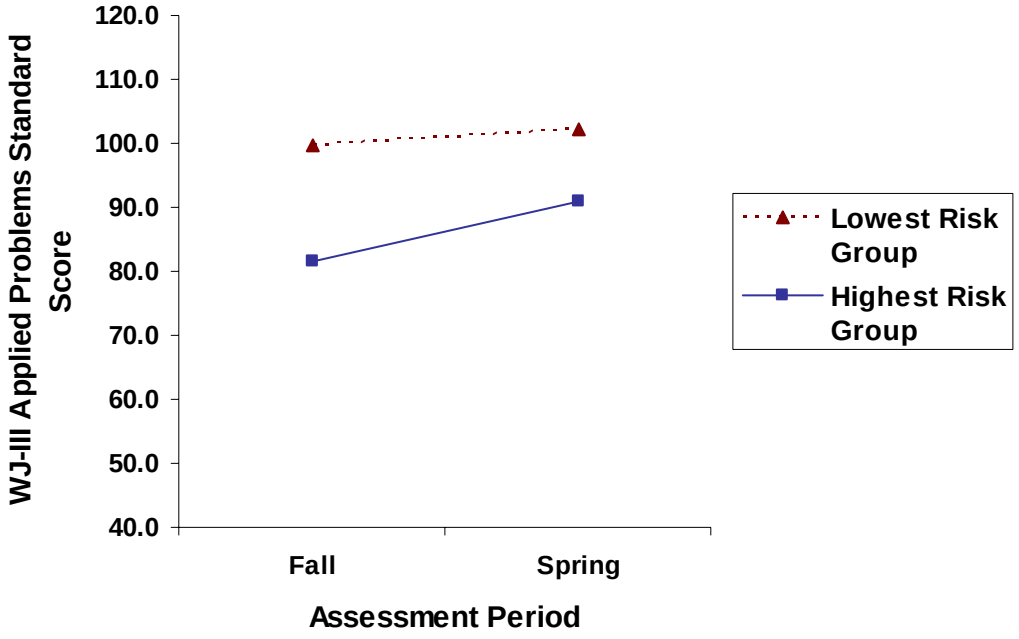


Figure 10. Growth in Color Knowledge (Color Naming Task) by Cumulative Risk (2005-2006 Cohort)

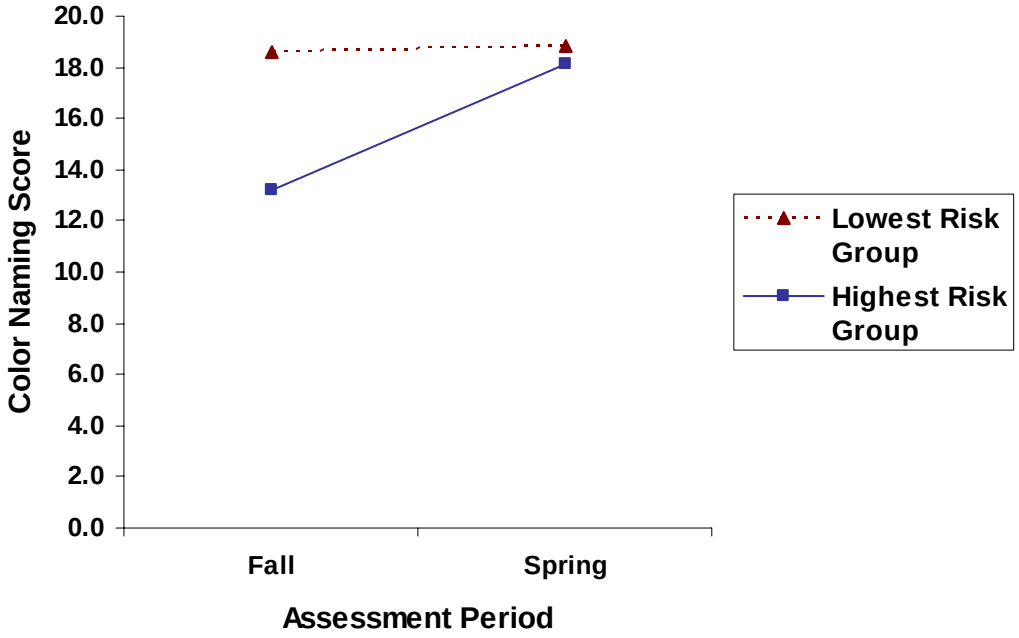
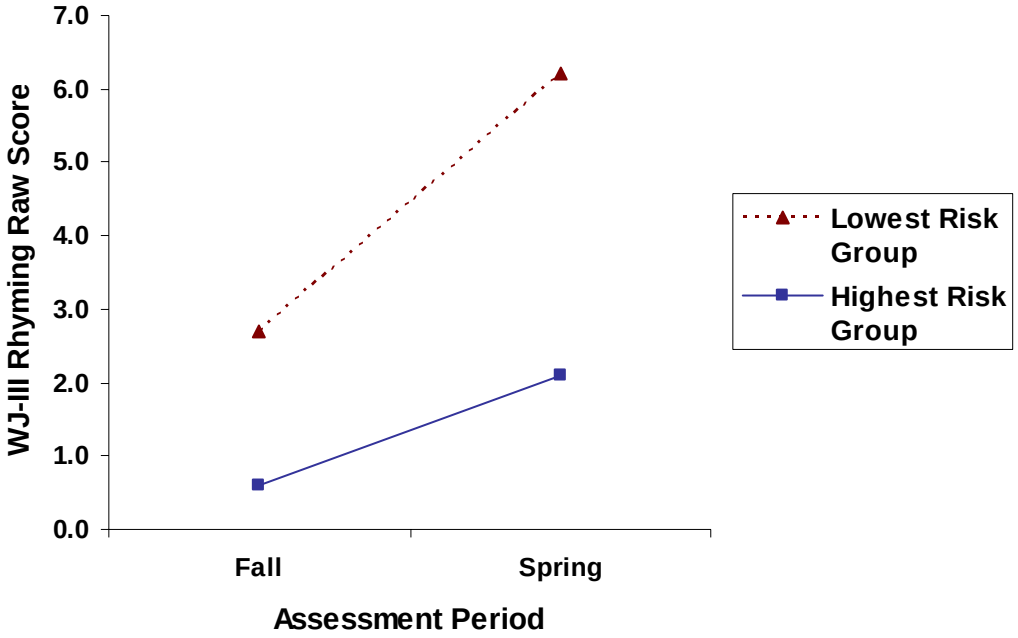


Figure 11. Growth in Phonological Awareness (WJ-III Rhyming) by Cumulative Risk (2005-2006 Cohort)



Differences in Child Outcomes by English Proficiency Levels

For the fourth series of analyses, we examined whether there were significant differences in the absolute level of performance or the rate of growth for children entering the program at different levels of English proficiency, based on individual assessments of oral language proficiency[12]. Children were categorized according to five proficiency levels ranging from Non-English speaker (1) to Limited English speaker (2-3) to Fluent English speaker (4-5). In addition, these analyses adjusted for children's risk factor level, gender, and attendance. To examine differences in the absolute level of scores by English proficiency level, separate analyses were conducted for the fall and spring scores. To examine differences in the amount of growth for children at different English proficiency levels, analyses of the change scores (spring scores minus fall scores) were conducted.

Differences in Level of Skills by English Proficiency Group

For children in the 2005-2006 cohort, as well as the 2003-2004 cohort, those at the lowest English proficiency levels (non-English speakers and/or those with limited English proficiency) tended to score lower in both the fall and spring than other children, especially those in the fluent groups (see Table 15). This pattern was evident for most measures of language/literacy skills (receptive language, rhyming skills, story concepts), math skills (applied problems, counting), and general knowledge (color knowledge, social awareness). For social skills, less proficient children scored lower in the fall but not the spring for the 2005-2006 sample, although this difference remained in the spring for the 2003-2004 sample. There was little or no difference in problem behaviors for both cohorts. Conversely, children at the highest English proficiency level (those fluent in English) scored higher in both the fall and spring than less proficient children on language/literacy skills (receptive language, rhyming, story concepts, naming letters) and math skills (applied problems, counting).[a]

Differences in Rate of Growth by English Proficiency Group

2005-2006 Cohort. These analyses examined whether there were differences in the amount of growth for children at different English proficiency levels for the 2005-2006 sample. Although their scores were lower, children at the lowest English proficiency level made greater gains over the More at Four year than more proficient children in several areas: Receptive language skills [$F(4, 122)=4.89$, $p<.01$] (see Figure 12); applied math skills [$F(4, 113)=7.02$, $p<.001$] (see Figure 13); color knowledge [$F(4, 125)=19.14$, $p<.001$] (see Figure 14); social awareness [$F(4, 123)=5.44$, $p<.001$] (see Figure 15); and social skills [$F(4, 99)=3.67$, $p<.01$] (see Figure 16). Note that these figures illustrate the differences between the performance of children at the lowest and highest English proficiency levels. Not surprisingly, in contrast, children at the highest proficiency level made greater gains than children at the lowest proficiency level on

[a] For both cohorts, we also examined whether children at lower levels of English proficiency were in classrooms with lower average quality or higher average proportions of More at Four children to determine whether there were differences in the quality of their experiences. We could not statistically test these differences because there were children at different proficiency levels within a classroom, and therefore, the measurement of English proficiency level categories and classroom characteristics were not independent. However, examination of the means for each proficiency group provided no evidence that the least proficient children were in lower-quality classrooms for either cohort; if anything, slightly the reverse was true (2003-2004 range=5.2-5.4 and 2004-2005 range=4.0-4.2). There was a slight trend indicating that the least proficient children were in classrooms with higher proportions of More at Four children, although these differences were slight (2003-2004 range=.86-.91 and 2005-2006 range=.83-.89).

rhyming [F(4, 124)=3.42, p<.05] (see Figure 17). For the remaining measures (story concepts, naming letters, counting, and problem behaviors), children exhibited similar rates of gain regardless of their initial English proficiency level.

Combined Cohort. These analyses tested differences in children's growth by English proficiency level for the combined cohort to examine whether these results were consistent with those for the 2005-2006 data alone. Children at lower levels of English proficiency showed greater developmental growth on measures of receptive language skills [F(4, 240)=6.69, p<.001], applied math skills [F(4, 219)=8.98, p<.001], color knowledge [F(4, 243)=36.13, p<.001], social awareness [F(4, 241)=9.63, p<.001], social skills [F(4, 226)=8.60, p<.001], and problem behaviors (decreases in problem behaviors) [F(4, 228)=3.44, p<.01]. In contrast, children at higher levels of English proficiency showed greater gains in rhyming skills [F(4, 231)=10.46, p<.001)]. There were no differences in the rate of growth over the pre-k year by English proficiency levels for other measures of language/literacy skills (story concepts and naming letters) or math skills (counting). These findings are consistent with the results of the 2005-2006 cohort described above, with the one exception of problem behaviors, which showed no differences in the single sample. In addition, there was one significant difference in the amount of gain exhibited by the two cohorts. The 2005-2006 cohort exhibited greater growth than the 2003-2004 cohort on receptive language [ß (se)=2.19 (.81), p<.03]. There were no cohort differences on the remaining measures. Further, total risk level was no longer a significant predictor of children's growth when individually-assessed English proficiency level was included in the analyses, although it is difficult to disentangle these two factors since nearly two-thirds (65%) of the children in the highest risk category (level 3+) were also in the lowest English proficiency group (level 1) across both cohorts.

Table 15. Child Outcomes by English Language Proficiency

Outcome	Language Proficiency Level[a]	2003-2004 N	2003-2004 Fall Mean (SD)	2003-2004 Spring Mean (SD)	2005-2006 N	2005-2006 Fall Mean (SD)	2005-2006 Spring Mean (SD)
Language and Literacy							
Receptive Language (PPVT-III[b])	**1**	74	60.4 (16.8)	68.0 (16.9)	104	55.9 (14.8)	67.2 (16.5)
	2	16	80.6 (8.2)	85.6 (11.2)	27	74.0 (10.7)	77.6 (15.5)
	3	70	82.4 (13.8)	89.5 (11.0)	56	83.5 (11.3)	88.2 (11.3)
	4	113	91.1 (11.1)	94.8 (9.5)	113	88.1 (13.5)	93.9 (12.9)
	5	147	98.5 (11.9)	99.8 (11.6)	122	98.3 (12.3)	102.9 (10.3)
	Significant group differences[c]:		1<2,3<4<5	1<2,3,4,5 2<4,5 3,4<5	Sig[c]:	1<2<3,4<5	1<2<3,4<5
Rhyming (WJ-III[d])	**1**	52	0.5 (0.9)	1.5 (1.5)	119	0.3 (0.7)	1.5 (1.9)
	2	16	0.8 (0.9)	2.1 (2.7)	27	0.9 (2.0)	2.3 (2.6)
	3	66	1.4 (2.2)	3.2 (3.4)	56	1.3 (1.9)	3.6 (3.8)
	4	109	1.7 (2.3)	4.6 (3.8)	111	2.0 (2.6)	4.3 (3.8)
	5	145	3.3 (3.5)	6.7 (4.4)	123	3.8 (3.8)	6.2 (4.0)
	Significant group differences[c]:		1,2,3,4<5	1,2,3,4<5 1<4	Sig[c]:	1,2,3,4<5 1<4	1,2,3,4<5 1<3,4

[a] These categories represent the fluency scores on the PreLAS 2000, an individual assessment of English language oral proficiency. Fluency level 1=non-English speaker, 2 & 3=limited English speaker, 4 & 5=fluent English speaker.
[b] Indicates standardized, norm-referenced measure with mean=100, SD=15.
[c] Significant comparisons reported represent significant group differences based on all possible pairwise comparisons of risk factor groups using adjusted p-values for multiple comparisons.
[d] Possible range=0-17.

Table 15. Child Outcomes by English Language Proficiency (continued)

Outcome	Language Proficiency Level	2003-2004			2005-2006		
		N	Fall	Spring	N	Fall	Spring
			Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)
Language and Literacy							
Story and Print Concepts[a]	**1**	63	1.3 (1.1)	3.1 (2.0)	120	1.2 (1.3)	3.1 (2.1)
	2	15	2.2 (1.8)	4.5 (2.6)	27	1.7 (1.4)	3.2 (2.1)
	3	69	2.2 (1.8)	4.0 (2.3)	54	2.3 (1.6)	3.7 (2.4)
	4	110	3.2 (1.8)	4.9 (2.2)	111	3.3 (2.2)	5.1 (2.3)
	5	147	4.4 (2.2)	6.5 (2.3)	122	4.8 (2.3)	6.6 (2.1)
	Significant group differences:		1,2,3,4<5 1<3<4	1,2,3,4<5 1,3<4	Sig:	1,2,3,4<5 1,2<4 1<3	1,2,3<4<5
Naming Letters[b]	**1**	76	2.9 (6.7)	10.3 (9.4)	121	1.8 (4.1)	10.8 (9.5)
	2	16	2.5 (3.6)	12.6 (8.5)	27	3.8 (7.0)	12.7 (10.8)
	3	70	5.3 (7.6)	13.8 (9.9)	56	4.2 (7.2)	13.2 (9.8)
	4	112	6.1 (7.3)	14.7 (9.5)	114	7.8 (8.7)	15.7 (9.2)
	5	147	9.8 (9.0)	19.0 (8.1)	123	11.5 (9.5)	20.8 (6.9)
	Significant group differences:		1<5 4<5	1<5 4<5	Sig:	1,2,3,4<5 1<4	1,2,3,4<5

[a] Possible range=0-14.
[b] Possible range=0-26.

Table 15. Child Outcomes by English Language Proficiency (continued)

Outcome	Language Proficiency Level	2003-2004			2005-2006		
		N	Fall	Spring	N	Fall	Spring
			Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)
Math							
Applied Problems (WJ-III[a])	**1**	41	79.3 (14.7)	87.8 (12.1)	71	76.1 (14.2)	89.5 (13.5)
	2	15	84.5 (10.6)	89.3 (8.2)	22	87.4 (10.2)	89.1 (12.5)
	3	68	86.9 (14.8)	91.3 (14.5)	55	87.2 (14.7)	90.5 (12.6)
	4	110	95.9 (11.6)	95.2 (11.6)	110	94.0 (12.0)	97.1 (10.4)
	5	144	100.0 (12.6)	101.4 (10.4)	120	100.9 (12.3)	103.4 (10.3)
	Significant group differences:		1<3,4<5 2,3<4<5	1,2,3,4<5 1<4	Sig:	1<2,3,4,5 2<5 3<4<5	1,2,3<4<5
Counting Task[b]	**1**	63	8.3 (6.4)	14.5 (10.1)	107	7.4 (4.0)	14.2 (8.5)
	2	14	7.9 (5.1)	12.9 (6.1)	26	8.9 (8.9)	13.6 (7.4)
	3	64	10.8 (6.7)	19.4 (12.2)	56	11.0 (7.5)	18.7 (9.4)
	4	111	11.9 (6.8)	18.9 (10.3)	113	12.7 (8.2)	19.7 (10.5)
	5	146	15.5 (9.9)	23.4 (11.7)	123	15.0 (8.2)	24.6 (10.8)
	Significant group differences:		1<5	1<5	Sig:	1<3,4,5 2<5	1<3,4,5 2,3,4<5

[a] Indicates standardized, norm-referenced measure with mean=100, SD=15.
[b] Possible range=0-40.

Table 15. Child Outcomes by English Language Proficiency (continued)

Outcome	Language Proficiency Level	2003-2004			2005-2006		
		N	Fall	Spring	N	Fall	Spring
			Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)
General Knowledge							
Color Knowledge[a]	**1**	76	10.6 (6.5)	17.3 (4.1)	121	11.0 (6.9)	17.4 (4.6)
	2	16	17.2 (5.1)	19.3 (2.0)	27	14.3 (5.7)	17.6 (4.9)
	3	70	15.8 (5.6)	18.3 (3.1)	56	16.3 (5.1)	18.8 (2.4)
	4	114	17.5 (4.4)	19.1 (2.3)	114	17.5 (4.2)	19.1 (2.3)
	5	149	18.9 (2.6)	19.5 (1.1)	123	18.9 (2.7)	19.5 (1.2)
	Significant group differences:		1<2,3,4,5 3<5	1<4,5	Sig:	1<2,3,4,5 2<5	1<3,4,5
Social Awareness[b]	**1**	77	1.5 (1.3)	3.1 (1.6)	119	1.3 (1.1)	2.7 (1.4)
	2	16	3.5 (1.6)	4.3 (1.4)	26	3.0 (1.5)	4.0 (1.5)
	3	70	3.6 (1.5)	4.6 (1.4)	56	3.8 (1.6)	4.6 (1.3)
	4	114	4.1 (1.4)	4.8 (1.2)	113	4.0 (1.5)	4.7 (1.2)
	5	149	4.7 (1.2)	5.0 (1.1)	123	4.7 (1.2)	5.1 (1.0)
	Significant group differences:		1<2,3,4,5 2,3,4<5	1<2,3,4,5	Sig:	1<2,3,4,5 2<4,5 3,4<5	1<2,3,4,5 2<4,5

[a] Possible range=0-20.
[b] Possible range=0-6.

Table 15. Child Outcomes by English Language Proficiency (continued)

Outcome	Language Proficiency Level	2003-2004			2005-2006		
		N	Fall	Spring	N	Fall	Spring
			Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)
Behavioral Skills							
Social Skills (SSRS[a])	**1**	81	93.9 (16.0)	105.9 (17.8)	94	98.1 (16.3)	112.2 (15.0)
	2	19	94.2 (15.4)	100.4 (17.2)	19	94.5 (15.6)	103.0 (16.8)
	3	73	100.9 (13.4)	107.0 (14.1)	45	100.1 (11.0)	106.8 (12.5)
	4	124	102.0 (14.7)	107.6 (15.0)	97	100.7 (14.0)	107.2 (14.2)
	5	156	105.6 (14.8)	110.7 (13.5)	103	108.0 (13.8)	113.7 (13.9)
	Significant group differences:		1<3,4,5 2<4,5 3<5	1,2<4,5	Sig:	1,2<5	NS
Problem Behaviors (SSRS[a])	**1**	82	101.2 (12.5)	98.6 (13.3)	98	96.2 (12.6)	94.4 (10.4)
	2	19	100.4 (12.7)	103.3 (14.0)	20	102.2 (15.7)	100.0 (13.1)
	3	71	98.0 (11.2)	98.0 (12.3)	48	98.7 (13.0)	99.3 (12.4)
	4	123	98.7 (12.6)	99.9 (13.0)	98	98.1 (12.7)	98.6 (11.7)
	5	159	96.7 (11.4)	99.0 (12.6)	106	96.6 (13.2)	96.8 (13.0)
	Significant group differences:		1>5	NS	Sig:	NS	NS

[a] Indicates standardized, norm-referenced measure with mean=100, SD=15.

Figure 12. Growth in Receptive Language Skills (PPVT-III) by English Proficiency (2005-2006 Cohort)

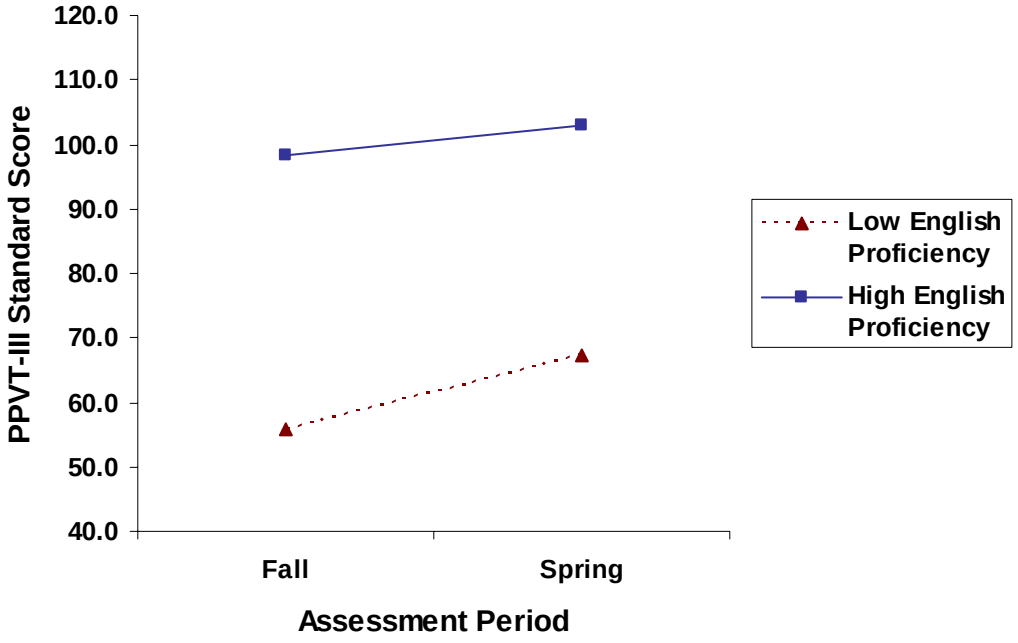


Figure 13. Growth in Math Skills (WJ-III Applied Problems) by English Proficiency (2005-2006 Cohort)

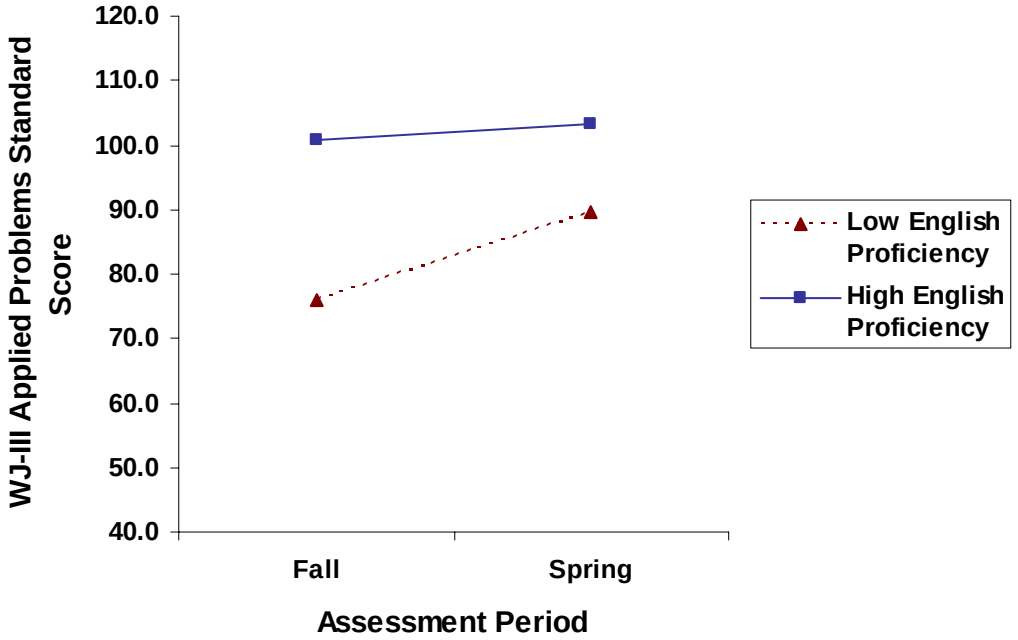


Figure 14. Growth in Color Knowledge (Color Naming Task) by English Proficiency (2005-2006 Cohort)

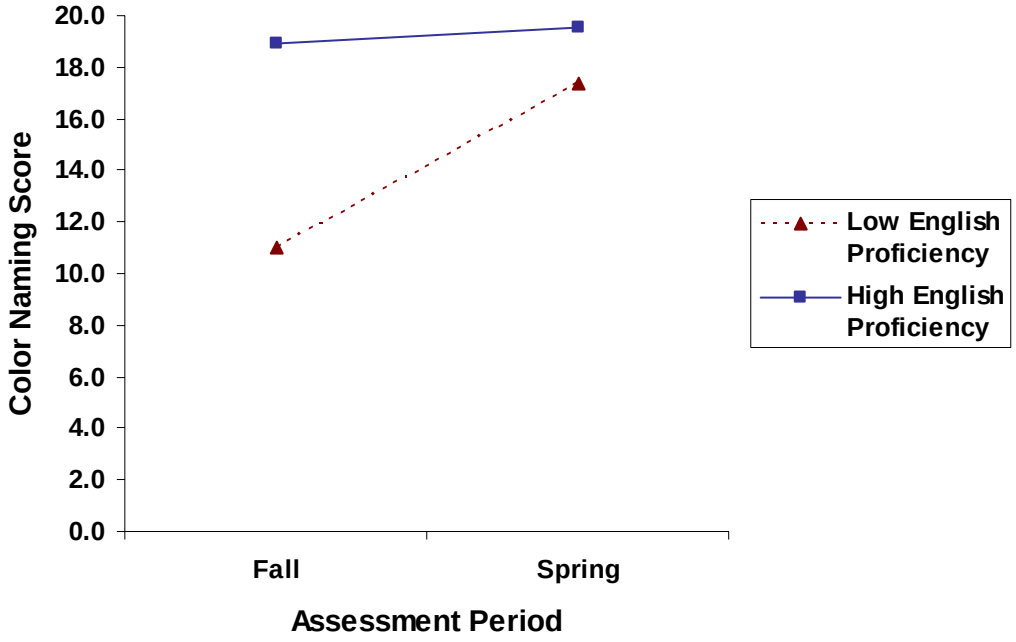


Figure 15. Growth in Social Knowledge (Social Awareness Task) by English Proficiency (2005-2006 Cohort)

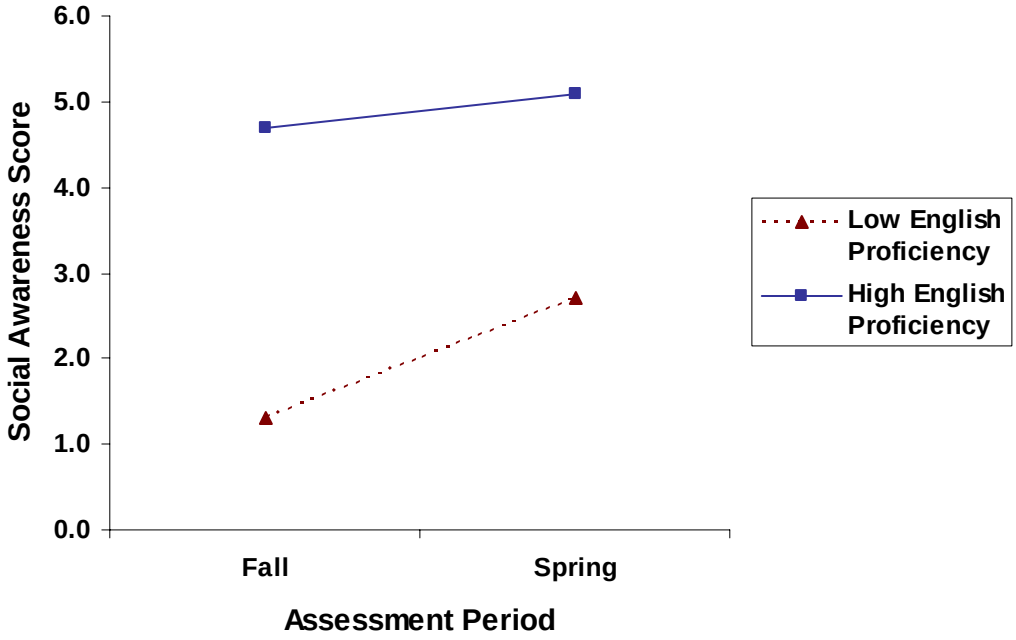


Figure 16. Growth in Social Skills (SSRS) by English Proficiency (2005-2006 Cohort)

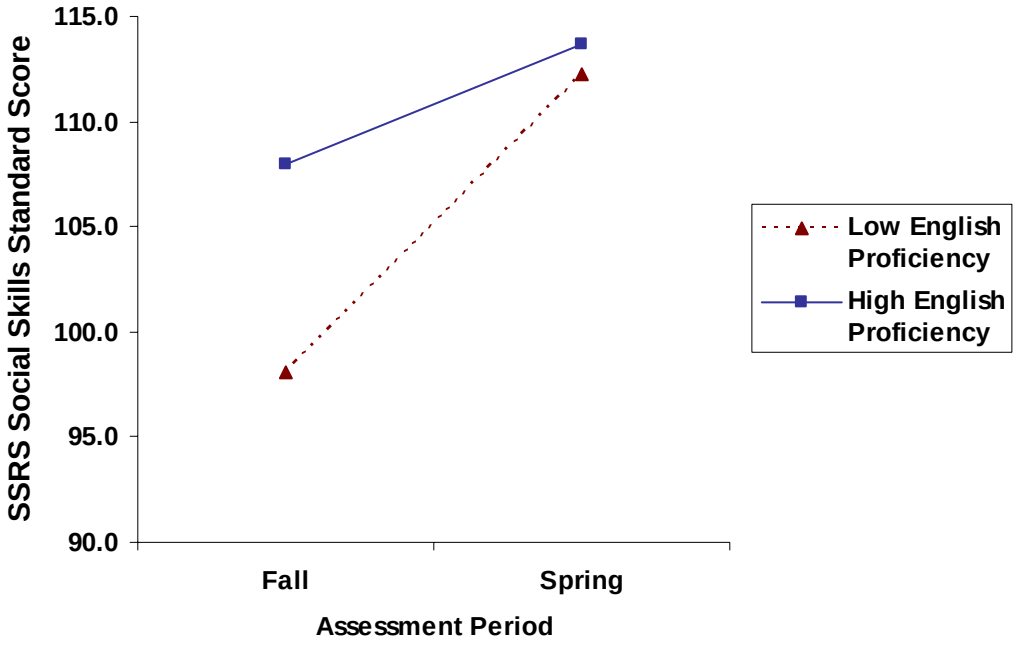
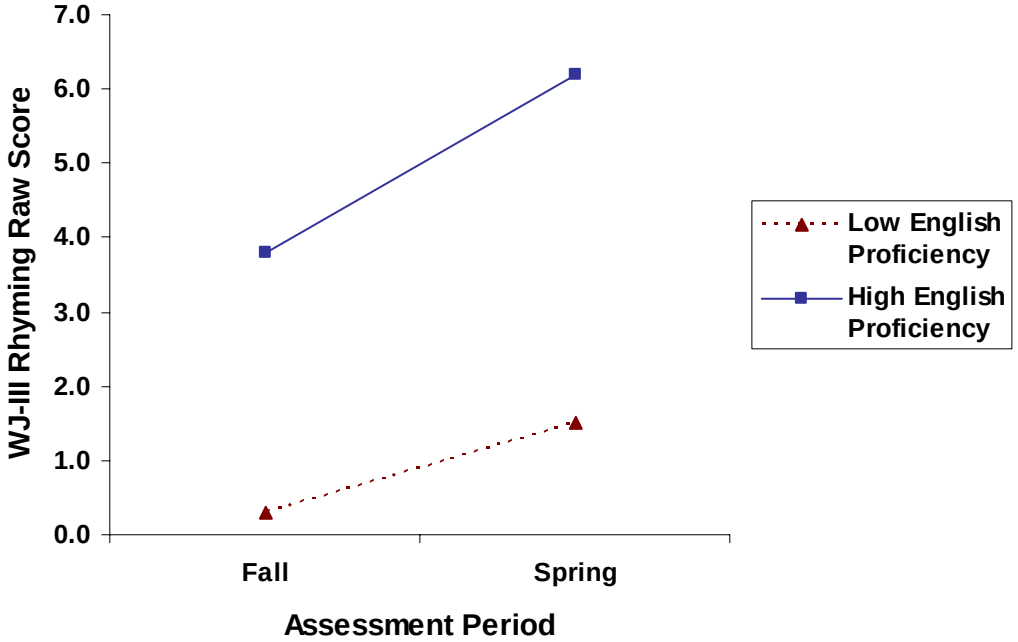


Figure 17. Growth in Phonological Awareness (WJ-III Rhyming) by English Proficiency (2005-2006 Cohort)



Subsample Analyses of Differences in Child Outcomes by Cumulative Risk Factor Levels

Given the previous findings, for the fifth series of analyses, we attempted to separate the effects of English proficiency from those of cumulative risk levels by excluding non-English speakers. We then examined whether there were significant differences in children's outcomes (level of skills or rate of growth) by cumulative risk level for this subsample of children. Because limited English proficiency is one of the individual risk factors and nearly two-thirds (65%) of the children in the highest risk category (level 3+) were also independently assessed as being in the lowest English proficiency group (level 1), it was not possible to statistically separate these two factors within the full sample. Therefore, we excluded children indicated as non-English speakers on the oral proficiency assessment (i.e., level 1) from these analyses. As in the previous analyses, children were categorized according to four levels of cumulative risk (0-3 from low risk to high risk).[a] To examine differences in the absolute level of scores by risk factor groups, separate analyses were conducted for the fall and spring scores. To examine differences in the amount of growth for children in different risk factor groups, analyses of the change scores (spring scores minus fall scores) were conducted. These analyses also adjusted for children's gender and attendance. In addition, given the reduced sample size, these analyses only examined the combined cohorts (2003-2004 and 2005-2006).

Subsample Differences in Level of Skills by Risk Factor Group

As seen in Table 16, children in the highest-risk group scored significantly lower than children in the other risk groups in both the fall and spring on nearly all the language/literacy (receptive language, rhyming, story concepts), math skills (applied problems, counting), and general knowledge (social awareness) measures for the combined cohort subsample excluding non-English speakers. On color knowledge, these differences were not significant by the spring, indicating that children in the highest-risk group had caught up to the other groups by the end of the pre-k year, with children in all groups scoring at or near the ceiling by the spring. In contrast, for letter knowledge, while the differences among risk groups were not significant in the fall, the highest-risk group was scoring significantly lower than the other groups in the spring. There were no differences among risk groups for measures of behavioral skills (social skills and problem behaviors). These findings indicate that, for the most part, children at greater risk are entering the pre-k program at lower levels and are remaining behind their peers in most skill areas by the end of the program year. These results for this subsample mirror the earlier results for the complete sample, suggesting that there are effects of cumulative risk apart from those of English proficiency on children's absolute skill levels.

[a] A total risk factor score was constructed based on More at Four eligibility guidelines, using income (eligibility for free lunch=2 points, reduced-price lunch=1 point, and full-price lunch=0 points) and additional risk factors (1 point each for limited English proficiency, identified disability, and chronic health condition). A four-level categorical variable was constructed, representing risk factor scores of 0, 1, 2, and 3-5.

Subsample Differences in Rate of Growth by Risk Factor Group

There was little difference in the rate of growth experienced by children at different levels of risk, as indicated by analyses of this combined cohort subsample excluding non-English speakers. For one measure, color knowledge, children in the highest-risk group made significantly greater gains than children in the other three groups [$F(3, 132)=4.13$, $p<.01$]. There were no differences on any of the other measures of children's language/literacy skills, math skills, general knowledge, or behavioral skills. These results contrast with the earlier results for the complete sample, where there were also greater gains for higher-risk children on receptive language, applied math skills, and social awareness, in addition to color knowledge. The discrepancies in these two sets of findings suggest that the differences in children's rates of growth (but not their absolute level of skills) are largely attributable to effects of English proficiency rather than to effects of cumulative risk.

Table 16. Child Outcomes by Risk Factor Levels for Subsample Excluding Non-English Speakers

Outcome	Risk Total Group[a]	Combined Cohort			
		Fall		Spring	
		N	Mean (SD)	N	Mean (SD)
Language and Literacy					
Receptive Language (PPVT-III[b])	**0**	62	97.9 (13.0)	60	99.2 (11.2)
	1	125	92.8 (13.2)	109	96.9 (12.9)
	2	488	90.4 (14.1)	435	95.3 (12.7)
	3+	72	82.9 (14.2)	63	88.6 (13.0)
	Significant group differences[c]:	3<0,1,2 2<0		3<0,1,2	
Rhyming (WJ-III[d])	**0**	60	3.5 (3.6)	59	6.5 (4.5)
	1	124	2.6 (3.2)	107	5.1 (4.2)
	2	488	2.2 (2.9)	431	4.9 (4.1)
	3+	72	1.0 (1.6)	61	2.7 (2.9)
	Significant group differences[c]:	3<0,1,2		3<0,1,2	

[a] A total risk factor score was constructed based on More at Four eligibility guidelines, using income (eligibility for free lunch=2 points, reduced-price lunch=1 point, and full-price lunch=0 points) and additional risk factors (1 point each for limited English proficiency, identified disability, and chronic health condition). A four-level categorical variable was constructed, representing risk factor scores of 0, 1, 2, and 3-5.
[b] Indicates standardized, norm-referenced measure with mean=100, SD=15.
[c] Significant comparisons reported represent significant group differences based on all possible pairwise comparisons of risk factor groups using adjusted p-values for multiple comparisons.
[d] Possible range=0-17.

Table 16. Child Outcomes by Risk Factor Levels for Subsample Excluding Non-English Speakers (continued)

Outcome	Risk Total Group	Combined Cohort			
		Fall		Spring	
		N	Mean (SD)	N	Mean (SD)
Language and Literacy					
Story and Print Concepts[a]	**0**	61	4.3 (2.4)	60	6.0 (2.4)
	1	123	3.8 (2.4)	109	5.8 (2.6)
	2	485	3.3 (2.2)	433	5.2 (2.4)
	3+	71	2.9 (2.1)	63	4.9 (2.7)
	Significant group differences:	3<0,1		3<0,1	
Naming Letters[b]	**0**	62	9.6 (9.0)	60	17.8 (8.4)
	1	124	8.9 (9.6)	109	17.6 (8.5)
	2	490	7.3 (8.4)	435	16.6 (9.3)
	3+	72	5.5 (7.7)	63	13.6 (9.7)
	Significant group differences:	NS		3<0,1,2	

[a] Possible range=0-14.
[b] Possible range=0-26.

Table 16. Child Outcomes by Risk Factor Levels for Subsample Excluding Non-English Speakers (continued)

Outcome	Risk Total Group	Combined Cohort			
		Fall		Spring	
		N	Mean (SD)	N	Mean (SD)
Math					
Applied Problems (WJ-III[a])	**0**	61	100.3 (13.7)	59	100.8 (10.9)
	1	121	99.0 (12.6)	109	100.1 (11.3)
	2	480	94.1 (13.6)	430	96.5 (12.4)
	3+	67	89.6 (13.9)	62	91.1 (15.3)
	Significant group differences:	3<0,1 2<0,1		3<0,1,2 2<0	
Counting Task[b]	**0**	60	14.9 (8.7)	60	23.1 (11.2)
	1	122	13.9 (7.7)	109	21.3 (11.6)
	2	481	12.7 (8.4)	433	20.5 (11.2)
	3+	71	9.8 (6.7)	63	17.0 (9.1)
	Significant group differences:	3<0,1		3<0,1,2	

[a] Indicates standardized, norm-referenced measure with mean=100, SD=15.
[b] Possible range=0-40.

Table 16. Child Outcomes by Risk Factor Levels for Subsample Excluding Non-English Speakers (continued)

Outcome	Risk Total Group	Combined Cohort			
		Fall		Spring	
		N	Mean (SD)	N	Mean (SD)
General Knowledge					
Color Knowledge[a]	**0**	63	18.8 (2.5)	60	19.3 (1.4)
	1	125	18.3 (3.4)	109	19.3 (1.7)
	2	492	17.7 (4.3)	436	19.1 (2.3)
	3+	72	15.3 (5.6)	63	18.8 (3.4)
	Significant group differences:	3<0,1,2		NS	
Social Awareness[b]	**0**	63	4.5 (1.3)	60	5.1 (1.0)
	1	125	4.3 (1.4)	108	5.0 (1.1)
	2	491	4.2 (1.4)	436	4.8 (1.2)
	3+	72	3.2 (1.5)	63	4.2 (1.3)
	Significant group differences:	3<0,1,2		3<0,1,2	

[a] Possible range=0-20.
[b] Possible range=0-6.

Table 16. Child Outcomes by Risk Factor Levels for Subsample Excluding Non-English Speakers (continued)

Outcome	Risk Total Group	Combined Cohort			
		Fall		Spring	
		N	Mean (SD)	N	Mean (SD)
Behavioral Skills					
Social Skills (SSRS[a])	**0**	58	105.3 (13.7)	53	110.8 (15.0)
	1	122	102.1 (15.9)	109	108.1 (14.0)
	2	481	101.5 (14.7)	421	108.5 (14.4)
	3+	68	102.9 (14.8)	59	109.4 (15.5)
	Significant group differences:	NS		NS	
Problem Behaviors (SSRS[a])	**0**	59	98.6 (12.0)	55	97.7 (12.6)
	1	122	97.6 (12.6)	111	99.3 (12.5)
	2	484	98.8 (12.7)	419	98.8 (12.5)
	3+	70	97.8 (12.8)	59	98.3 (13.2)
	Significant group differences:	NS		NS	

[a] Indicates standardized, norm-referenced measure with mean=100, SD=15.

Analysis Strategies

Changes Over Time

To investigate whether significant levels of growth occurred in the child outcomes assessed in English, we used the change scores (i.e., the difference between the spring and fall scores on each outcome) as the dependent variable. We conducted separate analyses for each outcome measure using a mixed models approach to account for multiple children clustered within each classroom[27]. Change scores were utilized because they exhibited a normal distribution within our sample, while, for several of the outcomes, the spring and fall scores were highly skewed.

Factors Affecting Level of Skills

We examined whether there were significant differences in the mean (or level) of the outcome measures across risk factor categories (0-3) and English proficiency levels (1-5). Post-hoc analyses were performed to test whether there were significant differences in the fall or spring scores by risk factor or by English proficiency levels. Analyses were conducted using a mixed models approach to adjust for nesting of children within classrooms. As a precaution against Type I error, the p-value was adjusted using the Tukey-Kramer correction for multiple comparisons. These analyses included the following covariates: gender, age at fall assessment, and the number of days the child attended the pre-k program prior to the fall assessment, as well as the number of days the child attended between the fall and spring assessments for the analyses of spring scores. For the analyses of English proficiency levels, risk factor was also included as a covariate.

Factors Affecting Rate of Growth

2005-2006 Cohort Analyses. To investigate the factors affecting the rate of change in the child outcomes assessed in English for the 2005-2006 cohort, we computed the change scores (i.e., the difference between the spring and fall scores) on each outcome measure and used the change score (i.e. gain or loss) as the dependent variable. We conducted separate analyses for each change score using a mixed model approach to account for clustering of children within classrooms[27]. This method examining change scores was chosen because the change score distribution for all the outcomes of interest appeared to be normally distributed, while the spring and fall scores were highly skewed for several of the outcomes.

To examine whether child or classroom characteristics differentially predicted change, we tested a series of hierarchical models. We first tested whether risk factor level was associated with the change score, then tested the addition of the child's English proficiency level as assessed in the fall, and finally, tested the addition of classroom quality in separate models for each of the three classroom quality measures (ECERS-R total, ELLCO classroom observation scale score, and CIS total). These analyses included the following covariates: gender, age at fall assessment, the number of days the child attended the pre-k program prior to the fall assessment, and the number of days the child attended between the fall and spring assessments. For the analyses of English proficiency level effects, risk factor was also included as a covariate, and for the analyses of classroom quality effects, both risk factor level and English proficiency level were included as covariates. We also tested whether the risk factor level by English proficiency level interaction was significant; in all cases, this interaction was nonsignificant and therefore was excluded from the final models.

Combined Cohort Analyses. General linear models were used to investigate children's growth over time for both cohorts, using individual change scores (i.e., the difference between spring and fall scores) for each child outcome measure as the dependent variable. This method examining change scores was chosen because the change score distribution for all the outcomes of interest appeared to be normally distributed, while the spring and fall scores were highly skewed for several of the outcomes.

Cohort (2003-04 or 2005-06) was included as a grouping variable in each analysis, and child gender, age at fall assessment, and total days of attendance during the program year were included as covariates. Differences in the rate of growth by risk factor level, assessed English proficiency level (adjusting for risk factor level), and classroom quality (adjusting for risk factor level and English proficiency level) were examined in a series of hierarchical models. In addition, the interactions of cohort with child characteristics (risk factor level and English proficiency) and with classroom quality were tested; in all cases, these interactions were nonsignificant and therefore were excluded from the final models.

Growth in Developmental Skills for Spanish Subsample

Two sets of analyses were conducted for the subsample of children administered measures in both English and Spanish in 2005-2006 (n=120). The purpose of these analyses was to examine whether children exhibited similar patterns of growth when assessed in English vs. Spanish and the extent to which changes in one language were related to changes in the other. The outcomes with both Spanish and English assessments included the measures of language/literacy skills (receptive language, rhyming, story concepts, letter naming), math skills (applied problems, counting), and general knowledge (color knowledge, social awareness). It is important to note that for the standardized measures (receptive language, rhyming, applied problems), the English and Spanish versions differed somewhat in content, so the absolute scores may not be directly comparable. For the remaining measures, the items on the English and Spanish versions were direct translations of one another.

Growth over the Pre-k Year

The first series of analyses examined the amount of growth the Spanish-speaking subsample of children exhibited on the various Spanish and English outcome measures over the More at Four program year. For these analyses, we tested whether the change scores (spring score minus fall score) were significantly different from zero. As shown in Table 17, children exhibited significant growth over the More at Four year on most of the English and Spanish measures, including language/literacy skills (rhyming, story concepts, letter naming), math skills (applied problems, counting), and general knowledge (color knowledge, social awareness). The one area that showed no significant growth in Spanish was receptive language skills, although children did show significant growth in English. Although children exhibited growth in both languages, the amount of growth varied across languages, with significantly more growth in English than in Spanish for all measures except story and print concepts (see last column of Table 17).

Table 17. Child Outcome Scores for Children with English and Spanish Assessments

Outcome	Domain	ENGLISH		SPANISH		Differences in English vs. Spanish Growth [a,c]
		Fall 2005 (n=74-120) **Mean** (SD) Range	**Spring 2006**[a,b] (n=89-108) **Mean** (SD) Range	**Fall 2005** (n=98-120) **Mean** (SD) Range	**Spring 2006**[a,b] (n=100-106) **Mean** (SD) Range	
Language and literacy	Receptive language[d] (PPVT-III, TVIP)	**55.3** (14.5) 23-88	**66.3***** (16.6) 32-107	**79.9** (14.8) 58-129	**79.6**[NS] (16.0) 55-122	***
	Rhyming[e] (WJ-III, Batería-III)	**0.4** (0.9) 0-7	**1.8***** (2.3) 0-12	**0.8** (1.1) 0-7	**1.2*** (2.0) 0-14	***
	Story and Print Concepts[f]	**1.2** (1.4) 0-6	**3.4***** (2.2) 0-9	**2.6** (2.0) 0-9	**4.3***** (2.4) 0-10	NS
	Naming Letters[g]	**1.4** (3.4) 0-22	**10.4***** (9.6) 0-26	**0.6** (1.7) 0-13	**1.2**** (2.9) 0-20	***
Math	Applied Problems[d] (WJ-III, Batería-III)	**77.8** (15.5) 47-115	**86.5***** (14.9) 51-116	**80.1** (15.6) 39-109	**84.0**** (15.8) 38-115	*
	Counting Task[h]	**7.6** (4.1) 1-22	**14.1***** (8.2) 2-40	**5.1** (3.6) 1-19	**8.1***** (5.1) 1-29	***
General knowledge	Social Awareness[i]	**1.4** (1.1) 0-5	**2.8***** (1.4) 1-6	**2.6** (1.3) 0-6	**3.0*** (1.4) 0-6	***
	Color Knowledge[j]	**10.9** (6.7) 0-20	**17.4***** (4.4) 1-20	**7.0** (5.5) 0-20	**9.3***** (5.5) 0-20	***

[a] *p < .05, **p < .01, ***p <.001, NS=nonsignificant.
[b] Significance levels indicate results of test of whether change scores are different from zero.
[c] Significance levels indicate result of test of whether the difference between the English Spring-Fall change scores and the Spanish Spring-Fall change scores was different from zero, with a correction for multiple comparisons.
[d] Indicates standardized, norm-referenced measures with mean=100, SD=15.
[e] Possible range for WJ-III=0-17; possible range for Batería-III=0-18.
[f] Possible range=0-14.
[g] Possible range=0-26.
[h] Possible range=0-40.
[i] Possible range=0-6.
[j] Possible range=0-20.

Associations between English and Spanish Growth

Given that the More at Four classrooms are primarily conducted in English, for the second set of analyses, we were interested in whether Spanish-speaking children's growth in skills in English was related to their level of skills in Spanish. Specifically, we were interested in whether children's level of initial skills and/or rate of growth when assessed in Spanish predicted their rate of growth when assessed in English. These analyses provided information about the extent to which children were exhibiting general patterns of skill development regardless of the language in which they were assessed vs. language-specific patterns of development. A series of regression analyses was conducted to test whether gains on the English measures (change scores) were related to children's initial scores in Spanish (fall scores) and/or gains on the Spanish measures (change scores) for the same outcomes (e.g., receptive language as measured by the PPVT-III and the TVIP).

As seen in Table 18, the results of these analyses indicated that both children's initial skills in Spanish and their rate of growth in Spanish, respectively, were positively associated with their rate of growth in English for several measures of language/literacy skills, math skills, and general knowledge: Story concepts [$F(1, 98)=8.04$, $p=.006$; $F(1, 98)=15.28$, $p<.001$]; applied math skills [$F(1, 58)=5.23$, $p=.026$; $F(1, 58)=4.97$, $p=.030$]; counting [$F(1, 79)=4.09$, $p=.047$; $F(1, 79)=4.53$, $p=.037$]; and social awareness [$F(1, 97)=6.70$, $p=.011$; $F(1, 97)=4.16$, $p=.044$]. For letter knowledge, children's English skill growth was positively related to their initial skill level in Spanish [$F(1, 102)=6.17$, $p=.015$], but not the rate of change in Spanish. In contrast, children's growth in Spanish, but not their initial skill level, was positively related to their growth in English for rhyming skills [$F(1, 98)=35.28$, $p<.001$]. Two measures (receptive language and color knowledge) showed no associations between growth in English and skills in Spanish.

Table 18. Associations of Growth on English Assessments with Initial Skills and Growth on Spanish Assessments

Domain	Assessment	Association with English Growth[a]	
		Initial Spanish Skill Level	Spanish Growth
Language and literacy	Receptive Language	NS	NS
	Rhyming	NS	***
	Story and Print Concepts	**	***
	Naming Letters	*	NS
Math	Applied Problems	*	*
	Counting Task	*	*
General knowledge	Social Awareness	*	*
	Color Knowledge	NS	NS

[a] *p < .05, **p < .01, ***p <.001, NS=nonsignificant.

Analysis Strategies

Growth over the Pre-k Year

To investigate whether significant levels of growth occurred in the child outcomes assessed in English and Spanish, we tested whether the change scores (i.e., the difference between the spring and fall scores on each outcome) for each measure were significantly different from zero. Change scores were utilized because they exhibited a normal distribution within our sample, while, for several of the outcomes, the spring and fall scores were highly skewed. For the comparisons of the amount of growth in English vs. Spanish, we tested whether the difference of the change scores (English change score minus Spanish change score) was significantly different from zero.

Associations between English and Spanish Growth

Change scores (i.e., difference between spring scores and fall scores) were computed for the English and Spanish outcomes for the subset of children assessed in both languages. A series of regression analyses were performed to assess whether changes in English outcomes were linearly associated with children's initial skill level in Spanish or their rate of growth in Spanish for the same measures.

Summary and Discussion

The More at Four Program has continued to grow through its fifth year of operations in 2005-2006, increasing from almost 11,000 children two years earlier to more than 17,000 children. While the program has grown substantially in size, most characteristics of the program have not changed noticeably over time. There continues to be a range of different types of sites and classrooms participating, with nearly half in public sites and half in community sites. The program has continued to serve a diverse group of at-risk children, with the majority in targeted groups such as very low-income families (i.e., children qualify for free lunch) and children who are unserved in another early education program at the time of enrollment.

The one area that has shown some change is the qualifications of teachers. One of the goals of the program is for lead teachers to have B-K licenses (or the equivalent), with classrooms given four years to reach this goal, according to the program guidelines. There has been a small increase in the level of teacher education over time, as indicated by an increased number with Bachelor's degrees or above (83% in 2005-2006 vs. 79% in 2003-2004) and a decrease in those with High School diplomas (2% in 2005-06 vs. 6% in 2003-04). There has been a somewhat larger increase in the number of lead teachers with B-K licenses (49% in 2005-2006 vs. 39% in 2003-2004), primarily attributable to teachers in public school settings (78% had B-K licenses in 2005-2006 vs. 66% in 2003-2004). In contrast, in community settings, there have been increases in the number of lead teachers with other early childhood credentials, such as CDA or NCECC (38% in 2005-2006 vs. 20% in 2003-2004).

In looking at the overall quality of the program, classroom practices were generally in the medium quality range based on ECERS-R observations of a random sample of classrooms operating in 2005-2006. The majority of classrooms scored in the medium quality range (86% scored 3.0-4.9), and nearly half (47%) scored at or above 4.5, the expected score based on the program guidelines. This represents a decrease in the overall quality of classroom practices compared to earlier years, although with one sample measured at a single point in time it is difficult to determine whether or not this is a general trend. As programs continue to scale up, it is always an issue of interest to examine whether quality is maintained. Given that this is the first year in the program's five-year history that a substantial decrease has been noted, it bears further investigation over time.

In contrast, observations indicated that teachers were fairly sensitive in their interactions with children, based on the CIS, suggesting that this is one area of strength in the program. Information about the quality of the classroom literacy environments using the ELLCO indicated that classrooms were doing a somewhat better job in terms of the organization of the literacy environment than the frequency of literacy-related activities. Moreover, scores tended to be higher for reading activities than writing activities, an area that might be important to consider for professional development efforts.

The results further indicated that there were no consistent predictors of classroom quality over time or across the various observational measures. Factors examined included teacher qualifications (B-K license or not), total class size, and classroom-level child characteristics (average proportion More at Four children, average risk level, average service priority level).

There was one significant finding for the 2005-2006 sample (but not the earlier sample), that having a lower proportion of More at Four children was associated with higher quality classroom practices (as measured by the ECERS-R). This finding suggests that efforts toward quality improvement may be even more critical for classrooms serving greater numbers of at-risk children. Similar results have been found in other studies of public pre-k programs, where having a lower proportion of poor children is associated with higher classroom quality[28]. There were no consistent associations between classroom quality and children's outcomes. For the 2005-2006 cohort only, higher quality was associated with greater increases in phonological awareness skills and social skills, but fewer increases in letter naming skills. The latter finding may be related to higher quality classrooms having less of a focus on teacher-directed instruction, which may be more closely associated with this type of learning, and/or a greater focus on more advanced language/literacy skills (such as phonological awareness). However, these findings were not maintained when the two cohorts were combined.

In terms of the effects on children participating in More at Four, children exhibited significant increases in developmental skills over the program year for all domains: Language/literacy skills (receptive language, rhyming, story and print concepts, naming letters), math skills (applied problems, counting task), general knowledge (color knowledge, social awareness), and social skills. There were no significant changes in children's problem behaviors, which remained slightly below the population mean (slightly fewer problem behaviors). These findings are consistent with the findings from previous years, and were maintained when the data from two cohorts (2003-2004 and 2005-2006) were combined.

There were some differences in skill levels based on children's risk at entry into the program, with those in the highest cumulative risk group or those at the lowest English proficiency level, entering with less developed skills. In some cases, these children made greater progress over the pre-k year. This was especially true for children entering with less proficient English skills, where they made greater gains on some skills across all areas of development—language/literacy skills, math skills, general knowledge, and social skills. However, it is also important to point out that for many domains, the skills of children at the highest risk level still remained substantially below those of their peers by the end of the pre-k year. There were some skills for which these differences were somewhat ameliorated by the end of the pre-k year, especially in the more recent sample—color knowledge for children with high risk totals; applied math skills and social skills for children with low English proficiency—indicating a success of the program.

Latino children comprise about one-fifth of the children served in More at Four, and this group is continuing to grow. Given the findings related to English proficiency, it was of particular interest to explore the growth of Spanish-speaking children on both English and Spanish measures. Even though the More at Four classrooms were primarily taught in English, these children exhibited significant growth in all domains on both the English and Spanish measures over the course of the pre-k year. The one exception was receptive language, where children showed significant growth in English but not in Spanish. Further, children's skill levels in Spanish were associated with their growth in English, suggesting that children's growth was not necessarily language-specific. Children who had higher initial skill levels in Spanish and/or demonstrated greater growth in Spanish skills also gained more in English skills during pre-k. These findings speak to the importance of promoting children's native language skills in conjunction with their growth in English skills.

In sum, these results show a similar pattern to findings from previous years, even as the program has continued to expand substantially each year. In contrast, while there were some positive increases in the level of teacher qualifications, the overall classroom quality scores were lower in the 2005-2006 sample than in earlier samples. Further, classroom quality, as measured across multiple areas (classroom practices, literacy environment, and teacher-child interactions), showed little association with children's outcomes. In terms of the overall effectiveness of the program for children, however, the results generally indicate a similar pattern to previous years. Children exhibited substantial developmental growth over the program year across multiple skill areas—language/literacy, math, general cognitive knowledge, and social skills. For Spanish-speaking children, growth occurred for skills assessed in both English and Spanish. While children at greatest risk, especially those at the lowest English proficiency levels, tended to make greater gains during pre-k, they entered the program with lower skill levels and still had not caught up to their peers in many areas by the end of the program year. These findings are consistent with a number of other large-scale studies that have found that pre-k participation was associated with sustained gains in language/literacy, math, and social skills as well as greater gains for children at greater risk.[29,30,31,32,33,34]

The More at Four Program is designed for children at risk, especially those who otherwise would not have been likely to participate in a pre-k program, thus providing them with opportunities for an early education experience. The evidence from the present year's evaluation, as well as those in the past, suggests that such experiences are beneficial and likely to help these children on the path to school success.

References

[1] For further details, see *More at Four Pre-kindergarten Program Guidelines and Requirements,* January 2005.

[2] Peisner-Feinberg, E. S., & Maris, C. L. (2005). *Evaluation of the North Carolina More at Four Pre-kindergarten Program: Year 3 Report (July 1, 2003-June 30, 2004)*. Chapel Hill, NC: FPG Child Development Institute, University of North Carolina at Chapel Hill.

[3] Peisner-Feinberg, E. S. (2003). *Child and Program Characteristics of the North Carolina More at Four Pre-kindergarten Program: Year 1 (January-June 2002).* Chapel Hill, NC: FPG Child Development Institute, University of North Carolina at Chapel Hill.

[4] Peisner-Feinberg, E. S., & Maris, C. L. (2005). *Evaluation of the North Carolina More at Four Pre-kindergarten Program: Year 2 (July 1, 2002-June 30, 2003).* Chapel Hill, NC: FPG Child Development Institute, University of North Carolina at Chapel Hill.

[5] Peisner-Feinberg, E. S., Elander, K. C., & Maris, C. L. (2006). *Evaluation of the North Carolina More at Four Pre-kindergarten Program: Year 4 Report (July 1, 2004-June 30, 2005) Program Characteristics and Services*. Chapel Hill, NC: FPG Child Development Institute, University of North Carolina at Chapel Hill.

[6] Peisner-Feinberg, E. S., & Maris, C. L. (2006). *Evaluation of the North Carolina More at Four Pre-kindergarten Program: Children's Longitudinal Outcomes and Classroom Quality in Kindergarten*. Chapel Hill, NC: FPG Child Development Institute, University of North Carolina at Chapel Hill.

[7] Kagan, S.L., Moore, E., & Bredekamp, S. (Eds.). (1995). *Reconsidering children's early development and learning: Toward common views and vocabulary. Goal 1 Technical Planning Group Report 95-03*. Washington, DC: National Education Goals Panel. See also http://govinfo.library.unt.edu/negp/ for a description of the National Education Goals.

[8] Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early Childhood Environment Rating Scale Revised Edition*. New York: Teachers College Press.

[9] Peisner-Feinberg, E. S., & Burchinal, M. R. (1997). Relations between child-care experiences and children's concurrent development: The Cost, Quality, and Outcomes Study. *Merrill-Palmer Quarterly, 43*, 451-477.

[10] Smith, M. W., & Dickinson, D. K. (2002). *Early Language Literacy Classroom Observation Toolkit, Research Edition*. Baltimore, MD: Brookes Publishing.

[11] Arnett, J. (1989). Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology, 10*, 541-552.

[12] Duncan, S. E., & De Avila, E. A. (1998). P*reLAS 2000*. Monterey, CA: CTB/McGraw-Hill.

[13] Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test Third Edition*. Circle Pines, Minnesota: American Guidance Service.

[14] Dunn, L. M., Padilla, E. R., Lugo, D. E., & Dunn, L. M. (1986). *Test de Vocabulario en Imagenes Peabody*. Circle Pines, MN: American Guidance Service.

[15] Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: The Riverside Publishing Company.

[16] Muñoz-Sandoval, A. F., Woodcock, R. W., McGrew, K.S., & Mather, N. (2005). *Batería III Pruebas de aprovechamiento*. Itasca, IL: Riverside Publishing.

[17] National Center for Early Development and Learning (2001). *Identifying Letters.* Unpublished instrument.

[18] FACES Research Team, modified from Story and Print Concepts tasks in: Jana M. Mason and Janice Stewart (1989). *The CAP Early Childhood Diagnostic Instrument*, prepublication edition, American Testronics.

[19] National Center for Early Development and Learning (2001). *Counting Numbers*. Unpublished instrument.

[20] FACES Research Team, modified from the Social and Communicative Competence tasks in: Jana M. Mason and Janice Stewart (1989). *The CAP Early Childhood Diagnostic Instrument* (prepublication edition), American Testronics.

[21] FACES Research Team, modified from the Color Concepts task in: J. M. Mason and J. Stewart (1989). *The CAP Early Childhood Diagnostic Instrument* (prepublication edition), American Testronics.

[22] Gresham, F., & Elliott, S. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.

[23] Dodge. D., & Colker, L. (1992). *The Creative Curriculum for Early Childhood Third Edition*. Washington, DC: Teaching Strategies Inc.

[24] Schickedanz, Judith A., Dickinson, David K., & Charlotte-Mecklenberg Schools. (2005). *Opening the World of Learning: A Comprehensive Early Literacy Program*. Parsippany, NJ: Pearson Early Learning Group.

[25] Smith, E. (2001). *Charlotte-Mecklenburg Schools Bright Beginnings Pre-Kindergarten Curriculum (Revised)*. Charlotte, NC: Charlotte-Mecklenburg Schools.

[26] 1st Edition: Hohmann, M., & Weikart, D. (1995). *Educating Young Children*. Ypsilanti, MI: High/Scope Press. 2nd Edition: Hohmann, M., & Weikart, D. (2002). *Educating Young Children Second Edition*. Ypsilanti, MI: High/Scope Press.

[27] Singer, J. D., (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 24,* 323-355.

[28] Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science, 9,* 144-159.

[29] Henry, G., Ponder, B., Rickman, D., Mashburn, A., Henderson, L., & Gordon, C. (2004). *An Evaluation of the Implementation of Georgia's Pre-K Program: Report of the Findings from the Georgia Early Childhood Study (2002-03).* Atlanta, GA: Georgia State University, Andrew Young School of Policy Studies.

[30] Magnuson, K.A., Meyers, M.K., Ruhm, C.J., & Waldfogel, J. (2004). Inequality in preschool education and school readiness. *American Educational Research Journal, 41,* 115-157.

[31] NICHD Early Child Care Network (2002). Early child care and children's development prior to school entry: Results from the NICHD Study of Child Care. *American Educational Research Journal, 39,* 133-164.

[32] NICHD Early Child Care Network (2005). Early child care and children's development in the primary grades: Follow-up results from the NICHD Study of Early Child Care. *American Educational Research Journal, 42,* 537-570.

[33] Peisner-Feinberg, E.S., Burchinal, M.R., Clifford, R.M., Culkin, M.L., Howes, C., Kagan, S.L., & Yazejian, N. (2001). The relation of preschool quality to children's cognitive and social developmental trajectories through second grade. *Child Development, 72,* 1534-1553.

[34] Zill, N., Resnick, G., Kim, K., O'Donnell, K., Sorongon, A., McKey, R., Pai-Samant, S., Clark,C., O'Brien, R., & D'Elio, A. (2003). *Head Start FACES 2000: A Whole-Child Perspective on Program Performance*. Washington, DC: Administration for Children and Families, US Department of Health and Human Services.